

Uniwersytet im. Adama Mickiewicza w Poznaniu  
Wydział Psychologii i Kognitywistyki

Szymon Bręński

**Kontrola procesu mówienia. Aspekty poznawczo-rozwojowe**

Control of the Speaking Process. Cognitive-Developmental Aspects

Rozprawa doktorska

przygotowana pod kierunkiem

dr hab. Stefana Frydrychowicza, prof. Uczelni



Poznań, 2024

## Streszczenie

**Cel.** Nadrzędnym celem pracy była analiza zmian rozwojowych w zakresie kontroli procesu mówienia i sprawdzenie, w jakim stopniu kontrola ta jest powiązana z funkcjami wykonawczymi. Wyróżniono trzy aspekty kontroli produkcji mowy, które biorą udział w powstawaniu i korygowaniu przejęzyczeń: (1) kontrola selekcji leksykalnej, (2) automonitorowanie mowy oraz (3) interupcje i reperacje. Podobnie, w projekcie uwzględniono trzy bazowe funkcje wykonawcze: hamowanie, odświeżanie pamięci i giętkość poznawczą. Podjęto próbę sformułowania odpowiedzi na pytania: Czy kontrola procesu mówienia zmienia się wraz z wiekiem od okresu późnego dzieciństwa do dorosłości? Jeśli tak, to (1) w jaki sposób osoby w odmiennych grupach wiekowych różnią się pod względem wybranych aspektów kontroli procesu mówienia oraz (2) jakie związki zachodzą między tymi aspektami kontroli procesu mówienia a funkcjami wykonawczymi?

### **Metoda.**

Badanie miało charakter quasi-eksperymentalny i zostało przeprowadzone w planie poprzecznym, obejmując osoby badane w trzech grupach wiekowych: dzieci w okresie późnego dzieciństwa ( $n = 51$ ), adolescentów ( $n = 50$ ) oraz młodych dorosłych ( $n = 50$ ). Uczestnicy wykonywali trzy główne zadania psycholingwistyczne: *Nazywanie obrazków* (paradygmat interferencji obrazka-wyrazu), *Monitorowanie sylab* (paradygmat monitorowania sylab), *Powtarzanie zdań* (paradygmat wywoływania reperacji). Pomiarowi podlegał sposób odpowiadania, poprawność odpowiedzi oraz czas reakcji. Ponadto, na potrzeby analizy związków między kontrolą produkcji mowy a kontrolą wykonawczą osoby badane wykonywały zadania: Zadanie z flankerami, n-wstecz, WCST. Zależności między funkcjami wykonawczymi a kontrolą procesu mówienia analizowane były w planie korelacyjnym.

**Wyniki.** Uzyskane rezultaty pokazują, że wraz z wiekiem rozwija się kontrola procesu mówienia. Zarówno skuteczność, jak i szybkość selekcji leksykalnej doskonalą się wraz z wiekiem. Potwierdza to obserwacja, że zewnętrzna dystrakcja semantyczna nie powoduje u osób dorosłych tak dużego opóźnienia w nazywaniu, jak u dzieci czy młodzieży, skutkując także mniejszą liczbą błędów i zakłóceń w postaci pauz wypełnionych. Również szybkość i dokładność monitorowania wewnętrznego rozwija się w okresie obejmującym badane grupy wiekowe, co może przekładać się na wykazaną w badaniu większą u osób dorosłych

skłonność do dokonywania reperacji ukrytych w strategii odroczenia w porównaniu do uczestników z młodszymi grupami wiekowymi. Wykazano także związki między kontrolą procesu mówienia, a funkcjami wykonawczymi, ale w większości przypadków były one słabe lub nieistotne, tylko częściowo potwierdzając postawione hipotezy.

**Dyskusja.** Badanie dostarczyło danych potwierdzających założenie o wzrastającym wraz z wiekiem poziomie kontroli mówienia w zakresie kontroli selekcji leksykalnej, monitorowania wewnętrznego oraz dokonywania reperacji. Jednocześnie obserwowane różnice między dziećmi, młodzieżą i dorosłymi mają raczej charakter ilościowy niż jakościowy. Z drugiej strony, wzrastająca wraz z wiekiem szybkość i dokładność w zakresie wybranych aspektów kontroli procesu mówienia może przekładać się na efekty jakościowe, skutkując większą płynnością mówienia, mniejszą liczbą zakłóceń, zawahań. Intrygującym wynikiem jest obserwowany wraz z wiekiem wzrost dokonywania reperacji wewnętrznych przez strategię odroczenia, co może skutkować mniejszą liczbą jawnych pomyłek i korekt w spontanicznej mowie. Wykazane związki między kontrolą procesu mówienia a funkcjami wykonawczymi pozwalają sądzić, że do pewnego stopnia ich rozwój jest sprzężony, jednak dalsze badania w tym temacie są konieczne.

**Słowa kluczowe:** proces mówienia, produkcja mowy, automonitorowanie, autoreperacje, nazywanie obrazków, funkcje wykonawcze, rozwój mowy

## SUMMARY

**Aim.** The main objective of this study was to analyze developmental changes in the control of the speaking process and to discover to what extent this control is related to executive functions. Based on the research literature, three aspects of speech production control have been identified that are involved in making and correcting slips of the tongue: (1) lexical selection control, (2) self-monitoring, and (3) interruptions and self-repairs. Likewise, the project included the three core executive functions: inhibition, updating, and switching, which were tested in relation to the control of the speaking process. An attempt was made to formulate answers to questions: Does the control of the speaking process change with age from late childhood to adulthood? If so, (1) how do people in different age groups differ in selected aspects of speaking control and (2) what are the relationships between these aspects of speaking control and executive functions?

### **Method.**

The study employed a quasi-experimental approach within a cross-sectional design, encompassing participants across three age categories: early school-aged children ( $n = 51$ ), adolescents ( $n = 50$ ), and young adults ( $n = 50$ ). The participants engaged in three main psycholinguistic tasks: Picture Naming (the picture-word interference paradigm), Syllable Monitoring (the syllable monitoring paradigm), and Sentence Repetition (the repair elicitation paradigm). The type of response, accuracy and reaction times, were measured. Furthermore, to analyze the connections between the control of the speech production and executive control, the participants performed the following tasks: the Flanker task, N-back, and WCST. The relationships between executive functions and the control of the speaking process were analyzed in a correlational design.

**Results.** The obtained results indicate that control of the speaking process develops with age. Both the efficiency and speed of lexical selection improve with age. This is confirmed by the observation that external semantic distraction does not cause as considerable a delay in naming in adults as in children or adolescents. It also results in fewer error rate and disruptions, such as filled pauses. Furthermore, the speed and accuracy of internal monitoring develop over the age groups studied. This development may contribute to a greater tendency among adults to make covert repairs using the postponement strategy, compared to younger participants. The study also revealed relationships between the control of the speaking process

and executive functions. However, these relationships were mostly weak or insignificant, only partially confirming the hypotheses

**Discussion.** The study provided data supporting the hypothesis that the level of speaking control increases with age, particularly in terms of lexical selection control, internal monitoring, and self-repairs. Nonetheless, the observed differences between children, adolescents, and adults are quantitative rather than qualitative. On the other hand, the speed and accuracy in selected aspects of speaking process control that come with age may contribute to qualitative effects, manifesting as increased speaking fluency, fewer disruptions, and hesitations. An intriguing finding is the observed age-related increase in internal repairs via the postponement strategy, potentially leading to fewer overt mistakes and corrections in spontaneous speech. The established connections between the control of the speaking process and executive functions suggest, to some extent, that their development is interlinked; however, further research on this topic is necessary.

**Key words:** speech production, self-monitoring, self-repairs, picture naming, executive functions, speech and language development



5.1.3	Interrupcja mówienia i sposoby dokonywania reperacji.....	166
5.1.4	Podsumowanie.....	171
5.2	Problemy badawcze i hipotezy.....	171
5.2.1	Pytania badawcze.....	172
5.2.2	Hipotezy.....	175
5.2.2.1	Grupa hipotez dotyczących pytań badawczych 1.1.1-1.3.2.....	175
5.2.2.2	Grupa hipotez dotyczących pytań badawczych 2.1.1-2.3.2.....	183
5.3	Struktura badanych zmiennych oraz operacjonalizacja.....	189
5.3.1	Definicje i operacjonalizacja zmiennych.....	189
5.3.2	Układ zmiennych.....	196
5.4	Metoda.....	198
5.4.1	Uczestnicy badania.....	198
5.4.1.1	Informacje wstępne.....	198
5.4.1.2	Procedura rekrutacji.....	198
5.4.1.3	Struktura demograficzna badanej próby.....	200
5.4.2	Procedura badania i aparatura.....	202
5.4.2.1	Zadania psycholingwistyczne.....	204
5.4.2.1.1	Nazywanie obrazków.....	204
5.4.2.1.2	Monitorowanie sylab.....	208
5.4.2.1.3	Powtarzanie zdań.....	214
5.4.2.2	Zadania mierzące kontrolę wykonawczą.....	220
5.4.2.2.1	BCST (WCST).....	220
5.4.2.2.2	Zadanie n-wstecz.....	223
5.4.2.2.3	Zadanie z flankerami.....	225
6	Wyniki.....	228
6.1	Wprowadzenie.....	228
6.2	Kontrola selekcji leksykalnej.....	228
6.2.1	Test hipotezy 1.1.1. Skuteczność selekcji leksykalnej wzrasta wraz z wiekiem..	231
6.2.2	Test hipotezy 1.1.2. Pod wpływem zewnętrznej dystrakcji semantycznej skuteczność selekcji leksykalnej obniża się u przedstawicieli wszystkich badanych grup wiekowych.....	232
6.2.3	Test hipotezy 1.1.3. Poziom odporności selekcji leksykalnej na zewnętrzną dystrakcję semantyczną wzrasta wraz z wiekiem.....	232
6.2.3.1	Dodatkowa analiza: Stosunek błędów mowy i braków odpowiedzi do ogólnej poprawności odpowiedzi.....	234
6.2.4	Test hipotez 1.1.4.-1.1.6. dotyczących rozwojowych oraz kontekstualnych uwarunkowań szybkości selekcji leksykalnej.....	235
6.2.4.1	Dodatkowa analiza (1): Procentowa wartość opóźnienia nazywania wywołana dystrakcją semantyczną.....	238
6.2.4.2	Dodatkowa analiza (2): Dystrybucja pauz wypełnionych w odpowiedziach uczestników badania.....	240
6.3	Efektywność automonitorowania wewnętrznego.....	241
6.3.1	Test hipotezy 1.2.1. Monitorowanie wewnętrzne zwiększa swoją dokładność wraz z wiekiem.....	246
6.3.2	Test hipotez 1.2.2.-1.2.4. dotyczących rozwojowych oraz językowych uwarunkowań szybkości monitorowania wewnętrznego.....	247
6.4	Reperacje błędów.....	253
6.4.1	Test hipotezy 1.3.1. Szybkość dokonywania interrupcji w błędnej wypowiedzi wzrasta wraz z wiekiem.....	255
6.4.1.1	Dodatkowa analiza (1): Porównanie procentowego rozkładu odpowiedzi z interrupcją w każdej z grup wiekowych.....	256

6.4.1.2 Dodatkowa analiza (2): Różnice w tempie mówienia między poszczególnymi grupami, a długość wypowiedzi z reparamentum.....	259
6.4.1.3 Dodatkowa analiza (3): Porównanie bezwzględnego czasu reakcji (przerwania wypowiedzi w celu wprowadzenia reperaturcji) między grupami wiekowymi.....	261
6.4.2 Test hipotezy 1.3.2. Częstość stosowania reperaturcji ukrytych (strategia odraczania) wzrasta wraz z wiekiem.....	262
6.4.2.1 Dodatkowa analiza (1): Stosowanie reperaturcji ukrytych w zadaniu <i>Powtarzania zdań</i> a tempo mówienia.....	264
6.4.2.2 Dodatkowa analiza (2): Spowolnienie tempa mówienia w wyniku stosowania strategii odraczania podczas reperaturcji ukrytych w różnych grupach wiekowych....	264
6.5 Kontrola procesu mówienia a poziom funkcji wykonawczych.....	266
6.5.1 Test hipotezy 2.1.1. Odporność selekcji leksykalnej na dystrakcję semantyczną zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w okresie od późnego dzieciństwa przez adolescencję do wcześniej dorosłości.....	269
6.5.2 Test hipotezy 2.1.2. Wielkość efektu interferencji semantycznej maleje wraz ze wzrostem poziomu funkcji wykonawczych u osób w okresie od późnego dzieciństwa przez adolescencję do wcześniej dorosłości.....	270
6.5.2.1 Dodatkowa analiza: <i>Nazywanie obrazków bez dystrakcji, a poziom funkcji wykonawczych</i> .....	271
6.5.3 Test hipotezy 2.2. Efektywność automonitorowania procesu mówienia wykazuje związek z poziomem funkcji wykonawczych.....	273
6.5.4 Test hipotezy 2.3. Sposób dokonywania reperaturcji wiąże się z poziomem funkcji wykonawczych.....	274
7 Dyskusja i wnioski.....	276
7.1 Dyskusja wyników.....	276
7.1.1 Kontrola selekcji leksykalnej.....	276
7.1.2 Automonitorowanie wewnętrzne.....	282
7.1.3 Reperaturcja błędów mowy.....	291
7.1.4 Kontrola procesu mówienia a poziom funkcji wykonawczych.....	297
7.2 Ograniczenia przeprowadzonego badania.....	305
7.3 Implikacje dla dalszych badań i praktyki.....	310
7.4 Podsumowanie.....	314
Bibliografia.....	318
Spis tabel.....	341
Spis rysunków.....	344
Załączniki.....	346



# Wprowadzenie

Ludzie należą do – jak to określa Deacon (1998) – „gatunku symbolicznego” (*symbolic species*), gdyż ludzka działalność przeniknięta jest używaniem symboli. Funkcja symboliczna, czyli wyjątkową umiejętność intencjonalnego zastępowania jednych obiektów innymi jest opanowywana przez dzieci na bardzo wczesnym etapie ontogenezy i wyraża się w zabawie, mowie, czy rysunku (Piaget i Inhelder, 1993). Jedną z najważniejszych cech symbolu jest to, że odnosi on do przedmiotów, zdarzeń, zjawisk i stanów rzeczy innych niż on sam, zachowując przy tym przejrzystość znaczeniową. Cecha ta przynależy także najbardziej wyjątkowemu z ludzkich systemów symbolicznych, czyli językowi. Mówienie, którego dotyczy niniejsza rozprawa, przez to że polega na generowaniu znaków języka, do pewnego stopnia także zachowuje transparentność. Oznacza to, że w pierwszej kolejności zwracamy uwagę na to, co mówi nadawca, czyli na treść komunikatu, a nie na to, w jaki sposób ten komunikat powstaje. Czynność mówienia pozostaje przez to niejako w tle swojego własnego wytworu, jakim jest wypowiedź (tekst przekazu). W pewnych sytuacjach jednak samo mówienie może stać się przedmiotem zainteresowania interlokutora, zwłaszcza jeśli ma to wpływ na jakość osiąganego porozumienia. Szczególnym przypadkiem takiej sytuacji jest popełnienie błędu mowy, np. w postaci przejęzyczenia, które – jeśli nie zostałyby skorygowane – mogłoby doprowadzić do niewłaściwego zrozumienia intencji nadawcy. Według szacunków podawanych przez Levelta (1999a) w normalnej, płynnej mowie wypowiadamy od dwóch do trzech słów na sekundę, myląc się przy tym średnio jedynie od jednego do dwóch razy na tysiąc wygenerowanych słów. Pytanie o to, jakie mechanizmy kontrolne zapewniają tak wysoką skuteczność w produkcji mowy, jest jedną z inspiracji dla niniejszej rozprawy. Kwestia ta jest tym bardziej interesująca, że – jak podają Blackmer i Mitton (1991) – w analizowanym przez nie materiale, reperacje wypowiedzi pojawiały się średnio częściej niż co 5 sekund. Oznacza to, że użytkownicy języka na bieżąco i w sposób (zapewne) ciągły monitorują proces zamiany myśli na słowa w mówieniu. Już samo to zjawisko wydaje się ciekawe i warte badania.

Zagadnienie produkcji mowy stanowi, obok badań nad odbiorem i rozumieniem mowy oraz akwizycją języka, jeden z głównych tematów podejmowanych przez psycholingwistykę (Bock, 1996; Dell, 1986). Jednocześnie, jak zauważa Harley (2001) istnieje pewna dysproporcja w zakresie liczby badań prowadzonych w ramach tych trzech głównych tematów. Według niego, badań nad produkcją mowy jest mniej niż nad rozumieniem, gdyż uchodzą one za trudniejsze do przeprowadzenia, głównie z powodu

ograniczeń w kontrolowaniu danych wejściowych (*input*) do zadań eksperymentalnych. Podobnie na gruncie polskim, psycholingwistyczne badania nad produkcją mowy stanowią mniejszość względem innych zagadnień. Jak można przeczytać w przedmowie do polskiego wydania podręcznika *Psycholingwistyka* (Berko Gleason i Ratner, 2005, s. 10) „modele wytwarzania mowy (...) nie były przedmiotem głębszego zainteresowania polskich psycholingwistów”<sup>1</sup>. Słowa te nie tylko oddają stan dotychczasowych badań, ale także stanowią zachętę do podjęcia wyzwania, czego efektem jest niniejsza rozprawa.

W badaniach produkcji mowy wyróżnić można dwie tradycje metodologiczne (Levelt, 1999a). Jedną z nich zajmuje się analizą spontanicznych przejęzyczeń, w celu odkrycia reguł rządzących procesami kodowania językowego. Drugą natomiast wykorzystuje podejście chronometryczne, aby na podstawie uzyskanych przez uczestników w zadaniach eksperymentalnych czasów reakcji oraz popełnionych przez nich błędów (główne zmienne w psychologii poznawczej) wnioskować o elementarnych procesach i stadiach przetwarzania języka, tworzących system produkcji mowy. Obie tradycje, choć w różnym stopniu, podejmują analizę błędów popełnianych przez użytkowników języka w spontanicznej lub motywowanej instrukcją eksperymentalną mowie. Jak zauważa Baars (1992) „podczas gdy większość ludzi spędza czas, próbując wszystko naprawić, psychologowie są przewrotnie oddani błędom”<sup>2</sup>. Podjęcie w niniejszej rozprawie tematu kontroli procesu mówienia, wychodzi naprzeciw tej konstatacji, gdyż na pierwszym planie stawia właśnie to, w jaki sposób ludzie próbują naprawić błędy w produkcji mowy lub nie dopuścić do ich wystąpienia. Tytuł pracy, oprócz odniesienia do badanego zjawiska, zdradza również podjęcie próby połączenia spojrzenia na kontrolę procesu mówienia z dwóch dopełniających się perspektyw. Pierwsza perspektywa, to teoretyczne zaplecze w postaci poznawczych modeli produkcji mowy, które aspirują do wyjaśnienia w jaki sposób myśl zamieniana jest na wypowiedź językową, przechodząc szereg transformacji, poczynając od konceptualizacji, a na artykulacji kończąc (Harley, 2001). Szczególne miejsce wśród tych modeli zajmuje teoria Levelta (1989; 1999a; 1999b; 2001b; Levelt i in., 1999, por. Frydrychowicz, 1999; Kurcz, 2005; 2011) oraz będąca jej częścią teoria podwójnej pętli percepcyjnej, wyjaśniająca działanie mechanizmów monitorowania procesu mówienia w oparciu o procesy odbioru mowy (Hartsuiker i Kolk, 2001; Hartsuiker, Kolk, i Martensen, 2005; Hartsuiker, 2007; Levelt, 1983; 1989; Levelt i in., 1999; Roelofs, 2005a; 2020a). Druga perspektywa, to spojrzenie na proces mówienia jako

---

1 Pełen cytat: „Modele wytwarzania mowy, omawiane w rozdziale 7., nie były przedmiotem głębszego zainteresowania polskich psycholingwistów. Wykorzystywano je jednak, zwłaszcza model Willema Levelta, do interpretacji wyników dotyczących rozumienia mowy, a w szczególności zdań dwuznacznych (por. prace Stefana Frydrychowicza)”.

2 Org. (ang.): „Whereas most humans spend their time trying to get things right, psychologists are perversely dedicated to error. „

element ludzkiego poznania i działania podlegający zmianom rozwojowym. Dane obserwacyjne dowodzą, że dzieci są w stanie skutecznie monitorować i kontrolować swoją mowę (przynajmniej pod względem treściowym) na bardzo wczesnym etapie (Clark, 1978), w zasadzie jak tylko zaczną posługiwać się mową (Hanley i in., 2016). Jednocześnie niewiele jest badań eksperymentalnie badających zmiany rozwojowe w zdolności monitorowania i korygowania mowy u dzieci i młodzieży. Przyjęty w niniejszej rozprawie plan poprzeczny i porównanie wybranych aspektów kontroli mówienia u dzieci w okresie późnego dzieciństwa, adolescentów i młodych dorosłych, ma na celu choć częściowe przybliżenie do zrozumienia na czym zmiany rozwojowe w monitorowaniu i kontroli mowy polegają. Tym, co spaja obie perspektywy (rozwojową i poznawczą) jest włączenie do analizy zagadnienia funkcji wykonawczych. Z jednej strony, od kilkunastu lat coraz mocniej akcentuje się konieczność odejścia od patrzenia na system produkcji mowy w kategoriach wyspecjalizowanych i autonomicznych modułów (Meyer i in., 2007), postulując związki monitorowania procesu mówienia z kontrolą wykonawczą (Nozari, 2018; Roelofs, 2020a; Shao i in., 2012; 2013). Z drugiej strony, badania w obszarze funkcji wykonawczych dowodzą ich intensywnego rozwoju w ciągu dzieciństwa, adolescencji i początku wczesnej dorosłości (Brzezińska i Nowotnik, 2012; Diamond, 2013; Józefacka-Szram, 2014). Uwzględnienie obu tych przesłanek, daje podstawę do poszukiwania wzajemnych zależności i rozpatrywania organizacji procesu produkcji mowy na tle rozwoju kontroli wykonawczej. Pomimo dostrzeganych przez autora ograniczeń wynikających z zastosowanych metod i planu badawczego, wyraża on nadzieję, że zaproponowane ujęcie rozszerza choć trochę dotychczasową wiedzę i pozwala spojrzeć na wąski problem kontroli procesu mówienia, w szerszym kontekście rozwoju poznawczego.

Pierwszy rozdział wprowadza terminologię i kontekst, głównie językoznawczych dociekań na temat mowy, języka i mówienia. Przedstawione zostały w nim ustalenia dotyczące struktury języka wypracowane przez klasyczne językoznawstwo strukturalistyczne, a także współczesne stanowiska należące do paradygmatu „mentalistyczno-kognitywistycznego” (Puppel, 2006, s. 6), które wchodzi w dialog z psychologią poznawczą. Po ustaleniu czynnościowego charakteru zjawiska, jakim jest mówienie, przedstawiona została w rozdziale drugim charakterystyka wewnętrznej strony (struktury) tego procesu, w postaci psycholingwistycznych modeli produkcji mowy. Prezentacja wybranych modeli teoretycznych stała się punktem wyjścia do bardziej szczegółowego rozważenia postulowanych przez nie mechanizmów kontroli. Temat ten podjęty został w rozdziale trzecim, łącząc perspektywę psycholingwistyczną z psychologicznymi badaniami nad

kontrolą wykonawczą. Rozdział czwarty dopełnia sygnalizowane przez tytuł wątki, przedstawiając wybrane aspekty rozwoju kontroli procesu mówienia na tle ogólnych zmian, dokonujących się w systemie produkcji mowy wraz z nabywaniem języka i rozwijaniem kompetencji językowej. Kolejne rozdziały przedstawiają badania własne. W rozdziale piątym dokonana została konceptualizacja i operacjonalizacja zmiennych, reprezentujących wybrane aspekty kontroli mówienia. W rozdziale tym opisane zostały także podjęte pytania badawcze i uzasadnienie postawionych hipotez, a także szczegółowo opisano metodę zbierania danych i sposób ich opracowania na potrzeby analizy statystycznej. Rozdział szósty przedstawia statystyki opisowe dla uzyskanych wyników oraz statystyczną weryfikację postawionych hipotez. W ostatnim rozdziale przeprowadzona została dyskusja wyników, uwzględniająca ograniczenia zastosowanej metody, ale także ich potencjał interpretacyjny i aplikacyjny.

Podczas pisania rozprawy autor starał się podążać za aktualnymi wytycznymi American Psychological Association, dotyczącymi standardów edytorskich naukowych tekstów psychologicznych (Skimina i in., 2022). Cytowania literatury przedmiotu i odnośniki bibliograficzne do niej, formatowanie tekstu oraz opis tabel, rysunków i wykresów przygotowywane były z intencją spełnienia standardu APA 7.

# 1 Mówienie na tle mowy i języka

## 1.1 Mowa i jej charakterystyka

Człowiek jedyny z istot żyjących obdarzony jest mową. Głos jest oznaką radości i bólu, dlatego posiadają go i inne istoty (rozwój ich posunął się bowiem tak daleko, że mają zdolność odczuwania bólu i radości, tudzież wyrażania tego między sobą). Ale mowa służy do określania tego, co pożyteczne czy szkodliwe, jak również tego, co sprawiedliwe czy też niesprawiedliwe. To bowiem jest właściwością człowieka, odróżniającą go od innych stworzeń żyjących, że on jedyny ma zdolność rozróżniania dobra i zła, sprawiedliwości i niesprawiedliwości i tym podobnych; wspólnota zaś takich istot staje się podstawą rodziny i państwa.

Arystoteles, *Polityka*<sup>3</sup>

Próba zdefiniowania czym jest mowa dotyka poniekąd samej istoty człowieczeństwa. To właśnie zdolność do posługiwania się mową była i nadal jest dla wielu tym, co odróżnia człowieka od innych zwierząt. Nie dziwi zatem, że wśród różnych określeń próbujących ukazać wyjątkowość człowieka takich jak *animal rationale* czy *animal cognitans*, podkreślających za Arystotelesem rozumność istot ludzkich, wymienia się także epitet *loquens* (Grzegorzczkova, 2007). Według Encyklopedii PWN *homo loquens* można tłumaczyć, jako „człowiek istota mówiąca” i wskazuje ono, że posługiwanie się mową można uznać za charakterystyczną cechę bytu ludzkiego.

Słownik terminologii językoznawczej definiuje mowę jako „zjawisko wzajemnego porozumiewania się ludzi głównie za pomocą dźwięków (wtórnie za pomocą gestów czy też w formie pisemnej)” (Gołąb, Heinz, Polański, 1970, s. 360). Podobnie, według Milewskiego (1967/2004, s. 5) mowę można zdefiniować jako „(...) porozumienie dźwiękowe dwóch osób, w którym jedna drugą o czymś powiadamia.” Z powyższych definicji, choć są one w pełni poprawne, nie wynika jeszcze, w jaki sposób mowa wyróżnia się spośród innych sposobów porozumiewania się u zwierząt. Z tego względu powyższe definicje dopełniane, czy też precyzowane są przez podział mowy na charakterystyczne tylko dla niej fazy. Pierwszą i podstawową fazą mowy jest mówienie. Podczas mówienia treść pojęciowa dostępna jedynie

---

3 Arystoteles (wyd. 2004).

nadawcy komunikatu jest zamieniana na kod językowy, którego fizycznym nośnikiem są fale dźwiękowe produkowane za pomocą ruchów aparatu artykulacyjnego. Posługiwanie się tym samym kodem przez odbiorcę komunikatu jest warunkiem koniecznym do wystąpienia rozumienia, czyli drugiej fazy mowy. Efektem mówienia i rozumienia jest tekst, czyli „utrwalone w pamięci lub na piśmie wypowiedzenie lub ich szereg” (Gołąb i in., 1970, s. 360). Tekst, jako twór konkretny o charakterze ciągłym do pewnego stopnia istnieje niezależnie od procesów mówienia i rozumienia, dlatego uznawany jest właśnie za osobną, trzecią fazę mowy (Milewski, 1967/2004; por. Polański, 2003). Ostatnią, ale nie mniej istotną fazą mowy jest język, rozumiany jako społeczny i abstrakcyjny twór mowy (Milewski, 1967/2004), „system społecznie wytworzonych znaków dźwiękowych (względnie wtórnie pisanych) oraz reguł określających ich użycie” (Gołąb, Heinz, Polański, 1970, s. 272).

Dalsze precyzowanie tego, czym jest mowa, wiąże się z szukaniem cech, które można by jej przypisać i uznać za wyróżniające na tle innych naturalnych sposobów komunikowania się. Hockett (1968, 1979) analizując różne formy komunikacji zwierząt i zestawiając je z ludzką mową wyróżnił 16 cech języka<sup>4</sup>, z których część jest współdzielona z innymi sposobami porozumiewania się obecnymi u innych gatunków. W żadnym jednak przypadku poza ludzką komunikacją, charakterystyki te – zdaniem Hocketta – nie są obecne jednocześnie w innym sposobie komunikacji, a niektóre zaś są obecne wyłącznie u ludzi.

1. Dwoistość (*duality of patterning*) – formy znaczące języka (znaki) składają się z elementów pozbawionych znaczenia, np. wyraz „kot” oznaczający czworonożnego ssaka, składa się z fonemów [k] + [o] + [t], które w izolacji nie posiadają żadnego znaczenia, natomiast połączone w innej kolejności mogą tworzyć wyraz o zupełnie odmiennym znaczeniu niż wyraz „kot” (np. tok, kto) lub wejść w skład innych wyrazów (np. oktawa, plotka).
2. Produktywność (*productivity*) lub inaczej otwartość (*openness*) – organizacja systemowa języka zapewnia nieograniczoną możliwość tworzenia nowych wypowiedzi.
3. Arbitralność (*arbitrariness*) – relacje między formą (dźwięk) i treścią znaku językowego (znaczenie) nie są wyznaczone relacją podobieństwa (brak naturalnego umotywowania) lecz są umowne.

---

4 Pierwotnie Hockett (1968) pisał o cechach języka (*features of language*), wymieniając 7 atrybutów typowo ludzkiego sposobu komunikowania się. Jednak jak zostanie dalej wykazane, tylko część spośród tych cech można przypisać językowi, a inne natomiast bardziej związane są z mówieniem.

4. Przemienność (*interchangeability*) – każda osoba może zarówno być nadawcą jak i odbiorcą, a porozumiewanie polega na wzajemnym i cyklicznym wymianianiu się przez rozmówców tymi rolami.
5. Specjalizacja (*specialisation*) – mowa służy wyłącznie komunikacji językowej i nie przynosi innych fizycznych (biologicznych) skutków poza tymi, które wynikają z faktu odebrania komunikatu przez odbiorcę.
6. Przemieszczanie się (*displacement*) – znaki języka są oderwane od „tu i teraz”, tzn. mogą odnosić się do przedmiotów i zjawisk nieobecnych w danej sytuacji komunikacyjnej (przemieszczenie czasowo-przestrzenne nazywane także odnośnością na dystans).
7. Transmisja kulturowa (*cultural transmission*) lub tradycyjność (*tradition*) – przyswajanie języka następuje w procesie międzypokoleniowego uczenia się (transmisja kulturowa), a nie poprzez przekaz genetyczny. Choć obserwuje się podobny proces przyswajania sposobów komunikacji u innych gatunków zwierząt na drodze uczenia się młodych od dorosłych osobników, to wyłącznie u ludzi obserwujemy intencjonalne nauczanie podejmowane przez dorosłych względem dzieci (Kurcz, 2005).
8. Kanał głosowo-słuchowy (*vocal-auditory channel*) – kanałem dystrybucji tekstu mówionego jest droga akustyczna. Narządy aparatu artykulacyjnego, powodując okresowe i lokalne zmiany ciśnienia powietrza, wytwarzają falę akustyczną, która jest odbierana jako dźwięki mowy (foniczność języka). Hockett (1979) podkreślał, że choć komunikacja wielu gatunków zwierząt oparta jest na kanale słuchowym, tylko u człowieka ma jednocześnie charakter głosowy.
9. Wielokierunkowa transmisja i jednokierunkowy odbiór (*broadcast transmission and directional reception*). Ta cecha mowy wynika z samej istoty dźwięku, który rozchodzi się koncentrycznie od źródła swojej emisji. Dzięki temu, wygenerowany przez nadawcę komunikat może zostać przyjęty przez odbiorcę, niezależnie od jego lokalizacji względem nadawcy, o ile znajduje się on w zasięgu fali akustycznej niosącej treść komunikatu. Jednocześnie odbiór komunikatu w danym momencie ograniczony jest do jednego źródła, co jest związane z szybkim zanikaniem sygnału językowego.
10. Szybkie zanikanie (*rapid fading*) – kodowanie komunikatu językowego w mówieniu odbywa się sekwencyjnie, co sprawia że sygnał językowy jest złożony z ciągu następujących po sobie dźwięków mowy. Przy tempie mówienia wynoszącym około

2-3 słów na sekundę, wyartykułowanych zostaje w tym czasie kilkanaście głosek. Aby złożony komunikat mógł zostać sprawnie odebrany, a jego wypowiedzianie zajęło stosunkowo niewiele czasu, konieczne jest szybkie następowanie po sobie segmentów fonetycznych. Poza tym z natury swojej obraz akustyczny jest dynamiczny i nietrwały.

11. Całkowite sprzężenie zwrotne (*total feedback/complete feedback*) – dzięki głosowo-słuchowemu kanałowi transmisji mowy nadawca może na bieżąco odbierać nadawany przez siebie komunikat, a tym samym dokonywać stałej ewaluacji jego treści. Daje to duże możliwości w zakresie kontroli czynności mówienia.
12. Semantyczność (*semanticity*) – znaki języka odnoszą się do rzeczywistości pozajęzykowej, tzn. mają swoje denotaty;
13. Dyskretność (*discreteness*) – komunikaty składają się z elementów, które wchodzą ze sobą w relacje opozycji, przez co różnią się od siebie w sposób nieciągły. Na przykład położenie języka w jamie ustnej w płaszczyźnie pionowej oraz poziomej, a także stopień otwarcia warg decydują o artykulacji głosek [a] i [ɔ]. Te drobne zmiany podczas artykulacji odbierane są jako odrębne fonemy, przez co powodują nieciągłą różnicę znaczeń między wyrazami, które różnią się zawartością właśnie tych głosek, np. las i los. Podobnie zmiana tylko jednej cechy dystynktywnej, jaką jest np. dźwięczność powoduje zmianę znaczenia w wyrazach „półka” i „bułka” (por. Polański, 2003).
14. Możliwość fałszu (*prevarication*) – język umożliwia nie tylko tworzenie komunikatów odnoszących się do rzeczywistości, ale też nieprawdziwych lub pozbawionych sensu (nadużywalność języka, por. Grzegorzczkowska, 2007).
15. Samozwrotność (*reflexiveness*) – opisu systemu języka można dokonać przy użyciu środków dostępnych w ramach tego systemu, to znaczy, że mówić można nie tylko o rzeczywistości pozajęzykowej (por. cecha semantyczności), ale także za pośrednictwem języka można wypowiedzieć się o nim samym.
16. Wyuczalność (*learnability*) – użytkownik danego języka (dialektu) może nauczyć się każdego innego. W praktyce oznacza to, że ludzie są w stanie porozumiewać się w więcej niż jednym języku i uczyć nowych już po okresie opanowania języka ojczystego.

Powyższa lista cech języka i mowy zaproponowana przez Hocketta (1960; 1968, 1979) powstawała stopniowo, obejmując początkowo siedem, a później trzynaście wymienionych cech i nie jest w pełni wyczerpującą charakterystyką. Według Żywiczyńskiego i Wacwicza (2015) propozycja Hocketta, pomimo upływu pół wieku od jej przedstawienia,

nadal ma duże znaczenie dla zrozumienia ludzkiej komunikacji i wciąż stanowi ważny punkt odniesienia dla współczesnych koncepcji lingwistycznych (por. Beecher, 2021). Nie oznacza to jednak, że to spojrzenie na mowę nie podlegało na przestrzeni tych lat krytyce. Lyons (1984) zauważa, że używanie klasyfikacji opartej na zestawie cech do porównania mowy z innymi systemami komunikacji nastrocza trudności, gdyż w przypadku porozumiewania się zwierząt trudno nieraz jednoznacznie uznać, że dana cecha występuje lub nie. Po pierwsze wynika to z nieprzystawalności różnych sposobów komunikowania się zwierząt i ludzi ze względu na ich różny stopień ewolucji, a tym samym różne drogi rozwoju porozumiewania się w filo- i ontogenezie. Po drugie, w przypadku ludzi jednym z istotnych kryteriów oceny porozumiewania się są intencje, jakie przyświecają rozmówcom. W przypadku oceny sygnałów używanych przez zwierzęta trudno jest przypisać im jednoznacznie intencje, na przykład nie da się z całą pewnością przypisać ofierze intencji wprowadzenia w błąd drapieżnika poprzez stosowanie jakichś mylących go sygnałów (Lyons, 1984). Z tego względu cecha „nadużywalności” zaproponowana przez Hocketta, przestaje być użyteczna w klasyfikacji sposobów porozumiewania się, gdyż jej występowanie w danym medium można stwierdzić, jedynie w odniesieniu do intencji komunikacyjnych nadawcy.

Podobne zarzuty formułują współcześnie Żywiczyński i Wacewicz (2015), wskazując że do pewnego stopnia cechy takie jak abstrakcyjność, przemieszczenie, czy nadużywalność charakteryzują bardziej zdolności poznawcze użytkowników języka niż sam kod komunikacji. Everett (2005) z kolei, odwołując się do swoich studiów nad językiem ludu Pirahã, wykazuje, że cechy przemienności, przemieszczenia i produktywności posiadają silne kulturowe uwarunkowania, co może powodować, że nie ujawniają się w pewnych dialektach na tyle, aby traktować je jako charakterystyczne, czy definicyjne dla mowy.

Kluczowym problemem zestawu cech zaproponowanych przez Hocketta, jest jednak to, że nie odnoszą się one wyłącznie do języka, ale także do jego specyficznego użycia w procesie mówienia. Na przykład Lyons (1984), omawiając cechy języka zaproponowane przez Hocketta, całkowicie pomija te, które dotyczą sposobu wydawania i odbierania dźwięku (kanał głosowo-słuchowy, wielokierunkowa transmisja i jednokierunkowy odbiór oraz szybkie zanikanie), a za najważniejsze uważa atrybuty, charakteryzujące język jako system znaków (por. Izert i Pachocińska, 1998). Problem odróżnienia języka od sposobu jego użycia uwidacznia się chociażby przy próbie określenia statusu języków migowych, które mimo iż nie spełniają wszystkich 16 cech definicyjnych Hocketa (1979), są w pełni równoprawnymi ekwiwalentami języków mówionych (Żywiczyński i Wacewicz, 2015). Podstawowa różnica dotyczy tego, że języki migowe nie korzystają z kanału głosowo-słuchowego, a z kanału

wizualno-przestrzennego (Dunaj, 2015). Pociąga to za sobą ograniczenie wielokierunkowej transmisji i otrzymywanie pełnego sprzężenia zwrotnego, jednocześnie dając szansę na spowolnienie a nawet okresowe zatrzymanie zanikania sygnału, co z kolei otwiera możliwość stosowania znaków w większym stopniu ikonicznych<sup>5</sup> niż w mowie głosowej. Z tego względu, jak zauważa Tomaszewski (2004) Polski Język Migowy (PJM) posiada wszystkie siedem cech wyróżnionych przez Hocketta (1968) w pierwotnym zestawieniu, które zwracały uwagę na wyjątkowość struktury języka, w oderwaniu od sposobu jego transmisji w procesie komunikacji. Zresztą dodanie kolejnych cech przez Hocketta (1960, 1979) związane było ze zwróceniem uwagi na wyjątkowość kanału komunikacji międzyludzkiej wynikającą ze specyficznej budowy i funkcji aparatu artykulacyjnego u ludzi<sup>6</sup>. Dzięki temu zabiegowi, rozbudowana o kolejne cechy klasyfikacja uwzględnia charakterystykę mowy we wszystkich jej czterech fazach. Jednocześnie tak ogólne zestawienie budzi kontrowersje, gdyż obejmuje we wspólną kategorię zjawiska, które tradycyjnie traktowane były jako odrębne.

Rozdzielanie zjawisk mowy na odrębne aspekty ma w językoznawstwie długą tradycję zapoczątkowaną poglądami szwajcarskiego badacza, prekursora europejskiego strukturalizmu, de Saussure'a (1991). Uważał on, że spośród ogółu zjawisk związanych z mową, którą określał terminem „langage” należy rozróżnić „langue”, czyli język i „parole”, czyli mówienie (Polański, 2003). Według de Saussure'a (1991) przez język należy rozumieć system znaków i reguł ich łączenia. Mówienie odnosi się natomiast do każdego jednostkowego aktu mowy – czynności w której użytkownik języka posługuje się nim, aby

---

5 Znaki ikoniczne opierają się na ich względnym podobieństwie do przedmiotu oznaczanego (Polański, 2003). W mowie opartej o kanał głosowo-słuchowy do znaków ikonicznych zaliczyć można wyrazy dźwiękonaśladowcze, które są do pewnego stopnia imitacją zjawiska, które oznaczają. W mowie opartej o kanał wzrokowo-przestrzenny potencjał ikoniczności jest znacznie większy (m.in. ze względu na trójwymiarowość kanału komunikacji) i obejmuje nie tylko leksykę (np. dłonie ułożone w kształt daszku oznaczają „dom”) ale i gramatykę, np. często liczba mnoga tworzona jest przez powielenie formy podstawowej znaku (Łozińska, Rutkowski, 2014). Analogiczny przykład można by jednak wskazać także dla języków realizowanych kanałem dźwiękowym, gdyż formy mnogie rzeczowników przekazują więcej materiału fonetycznego (są dłuższe) niż formy w liczbie pojedynczej, co nazywane jest ikonicznością diagramatyczną. Z tego względu językoznawcy kognitywni są skłonni przypisywać ikoniczności większą rolę w języku (niezależnie od kanału komunikacji) niż tradycyjne podejście strukturalistyczne (Taylor, 2007).

6 W mowie udział bierze nie tylko aparat artykulacyjny, ale także aparat odbiorczy związany ze zmysłem słuchu (tzw. analizator słuchowy). Jest on zaangażowany w odbiór mowy nie tylko po stronie odbiorcy w fazie rozumienia, ale także – po stronie nadawcy – ma istotny udział w fazie mówienia, pozwalając na kontrolę generowanego tekstu. Jednocześnie o specyfice ludzkiej mowy w większym stopniu decyduje budowa i funkcja aparatu artykulacyjnego niespotykana w takiej postaci u innych gatunków zwierząt poza człowiekiem. Żadne ze znanych gatunków zwierząt nie potrafi artykułować dźwięków mowy na ludzki sposób, nie licząc anegdotycznych wzmianek o mówieniu przez szympansy oraz imitacji głosu ludzkiego przez niektóre gatunki ptaków. Wiele gatunków natomiast wydaje się rozumieć ludzką mowę. W opisach przypadków uczenia szympansov i goryli języka migowego, wprost mówi się, że zwierzęta te „rozumiały” wiele słów języka mówionego (Kurcz, 2005). Choć trudno wnioskować, że inne gatunki rozumieją mowę tak samo jak ludzie, to jednak nie można zaprzeczyć temu, że potrafią ją odbierać i adekwatnie na nią reagować.

porozumieć się z innym użytkownikiem języka. Inaczej mówiąc, język jest abstrakcyjnym i społecznym aspektem mowy, będącym właściwym przedmiotem zainteresowania językoznawstwa w rozumieniu de Saussure'a, natomiast mówienie stanowi indywidualną realizację języka.

Obrazowo to rozróżnienie można przyrównać do gry w szachy (Polański, 2003). Mówienie, podobnie jak czynność grania w szachy wykorzystuje pewien określony zbiór elementów i reguł ich stosowania, który istnieje niezależnie od jednostkowej rozgrywki. Analogicznie, język, jako system znaków i reguł ich łączenia, istnieje niezależnie od jednostkowej wypowiedzi, a każdy akt mówienia jest konkretną realizacją abstrakcyjnego języka. Okoliczności w jakich przebiega gra w szachy, choć mogą determinować przebieg rozgrywki w żaden sposób nie wpływają na strukturę reguł, które są niezależne od warunków zewnętrznych. Tak samo język bytuje niezależnie od mówienia, które samo w sobie może przebiegać w różnych warunkach i być przez te warunki kształtowane.

De Saussure (1991, s. 41) motywował wprowadzoną przez siebie dychotomię twierdzeniem, że „oddzielając język od mówienia, oddzielamy tym samym: a) to, co społeczne, od tego, co indywidualne; b) to, co istotne, od tego, co uboczne i mniej lub bardziej przypadkowe”. Według MacNeilage'a (2005) rozróżnienie na abstrakcyjną językową formę (system języka „la langue”) i fonetyczną substancję (obserwowalna mowa jednostkowa „la parole”) odzwierciedla klasyczny platoński podział na idee ogólne i rzeczy jednostkowe. Z tego też względu, mając umocowanie w fundamentalnych dla zachodniej filozofii kategoriach pojęciowych, myśl de Saussure'a wywarła silny wpływ na późniejsze badania nad porozumiewaniem się ludzi. Przed wszystkim sformułowaniem przez de Saussure'a podziałem na mówienie i język wprowadzony został podział językoznawstwa na wewnętrzne i zewnętrzne (Izert i Pachocińska, 1998). Przedmiotem zainteresowania tego pierwszego miała być abstrakcyjna struktura języka, a zagadnienia związane z mówieniem w zasadzie zniknęły z horyzontu badań językoznawców, pozostając jednocześnie w obrębie zainteresowań innych dyscyplin (językoznawstwo zewnętrzne). Dychotomia ta nie pozostała także bez znaczenia dla psycholingwistycznych badań nad procesem mówienia i formułowanych na ich podstawie teoretycznych modeli tego procesu, które zaczęły powstawać w drugiej połowie XX wieku w celu wyjaśnienia w jaki sposób język realizuje się w mówieniu (zob. rozdział 2.).

W dalszej części tego rozdziału język i mówienie scharakteryzowane zostaną osobno, zgodnie z podziałem zaproponowanym przez de Saussure'a i wyznaczającym przedmiot badań językoznawstwa i innych nauk zajmujących się komunikacją ludzi. Ze względu na

tematykę niniejszej rozprawy pozostałe dwie fazy mowy, czyli rozumienie i tekst nie będą dalej szczegółowo opisywane<sup>7</sup>.

## 1.2 Język i jego różne ujęcia

Językoznawstwo wewnętrzne zaproponowane przez de Saussure'a miało za zadanie wypreparować z „niejednorodnej” i „niepoznawalnej” mowy zjawiska językowe, które następnie miało badać w oderwaniu od ich konkretnego, jednostkowego użycia, czyli mówienia. Zapoczątkowany przez tego badacza nurt językoznawstwa nazywany strukturalizmem koncentrował się na ujęciu języka jako systemu, interpretując wszystkie jego elementy w odniesieniu do innych (Polański, 2003). Wedle definicji samego de Saussure'a (1991, s. 36-37). język jest systemem „odrębnych znaków odpowiadających odrębnym wyobrażeniom” i jako „społeczny wytwór zdolności mowy” ma charakter umowny. Oznacza to, że znaki języka wchodzą ze sobą nawzajem w relacje mające charakter systemowy, natomiast z rzeczywistością pozajęzykową (światem pojęć i przedmiotów) wchodzą w relacje o charakterze umownym, motywowanym konwencją społeczną, a nie związkiem naturalnym. Podobnie określa język współczesna definicja zaproponowana przez Grzegorzyczkową (2007, s. 13), według której język to „system znaków konwencjonalnych, fonicznych<sup>8</sup>, służący do porozumiewania się o wszystkim (uniwersalny), charakteryzujący się tzw. dwuklasowością, tzn. udziałem gramatyki, umożliwiającej nieograniczone tworzenie nowych konstrukcji”. Definicja ta, zostanie rozwinięta poniżej.

### 1.2.1 Język w ujęciu strukturalistycznym

Tradycyjnie system języka podzielić można na trzy płaszczyzny: fonologiczną, semantyczną i syntaktyczną (Kurcz, 1987). Naczelną zasadą organizującą relacje między

7 Podział mowy na cztery fazy jest wtórny wobec dychotomii de Saussure'a. Dla językoznawców zarówno mówienie, jak i rozumienie znalazły się poza zakresem ich zainteresowania (językoznawstwo zewnętrzne). Z kolei badania psycholingwistyczne, bazujące na metodologii psychologii poznawczej intensywnie eksplorowały temat odbioru i rozumienia mowy, marginalizując w dużym stopniu badania procesu mówienia (Fromkin i Ratner, 2005; Harley, 2001; MacNeilage, 2008). W związku z tym wypracowano także odrębne paradygmaty badawcze dla badania odbioru i nadawania mowy. Rozumienie mowy jest więc traktowane w literaturze psycholingwistycznej jako odrębne zagadnienie i jako takie nie będzie szczegółowo omawiane w niniejszej rozprawie. Jednocześnie zagadnienie kontroli procesu mówienia nierozzerwalnie wiąże się z problemem oceny wygenerowanego tekstu pod względem jego poprawności językowej, jak i zrozumiałości dla odbiorcy oraz dla samego nadawcy. Z tego względu procesy związane z odbiorem i rozumieniem mowy zostaną uwzględnione w rozdziale 3., który poświęcony jest kontroli procesu mówienia.

8 De Saussure (1961) uważał, że foniczność języka związana z używaniem aparatu głosowego nie jest konieczna dla posługiwania się językiem i ludzie jako narzędzia języka mogliby używać gestów lub obrazów wizualnych zamiast akustycznych. Według niego „istota języka (...) jest obca dźwiękowemu charakterowi znaku językowego” (de Saussure, 1961, s. 24).

jednostkami poszczególnych podsystemów jest reguła opozycji. Oznacza to, że elementy podstawowe dla danej płaszczyzny są od siebie odróżnialne dzięki posiadaniu przeciwstawnych sobie cech, nie stanowią zaś ciągłego kontinuum. W ten sposób realizuje się cecha dyskretności sformułowana przez Hocketta (1968). Jednostkami podsystemu fonologicznego są fonemy, czyli abstrakcyjne reprezentacje dźwięków mowy. Ze względu na charakterystyczne dla każdego fonemu wiązki cech dystynktywnych, jednostki te wchodzą ze sobą w układy wzajemnych opozycji np. „dźwięczność-bezdźwięczność”. Jest to przejaw ogólniejszej zasady organizującej wzajemne relacje elementów systemu językowego na zasadzie przeciwieństw, np. forma i treść, jednostka i kategoria (Kurcz, 1987; Gołąb i in., 1970). Fonemy przynależące do formy języka, dzięki regułom fonotaktycznym mogą tworzyć morfemy będące nośnikami treści. Z kolei „morfemy stanowią składnik semantyczny języka, w którym reguły sensu decydują o wzajemnych relacjach morfemów leksykalnych, a reguły oznaczania stanowią o przedmiotach odniesienia tych morfemów” (Kurcz, 1987, s. 37). Trzecia płaszczyzna języka, związana jest z regułami syntaktycznymi, które odpowiadają za łączenie morfemów leksykalnych i gramatycznych w zdania.

Relacje między elementami systemu języka organizowane są również wedle zasady „całość-część”. Systemowość języka polega więc także na tym, że z prostych elementów składowych można zbudować struktury mające inne właściwości niż tylko suma tych elementów, gdyż między elementami zachodzą przekształcenia powodujące, że „z jednej formy ustrukturalizowania elementów możliwe są przejścia do innych form strukturalizacji” (Kurcz, 1987, s. 13). Zatem to, co jest pewną ustrukturowaną całością, jest jednocześnie strukturujące dla elementów prostszych (Frydrychowicz, 1999). Zasada ta realizuje się we wspomianej już za Hockettem (1968) dwustopniowości struktury języka, nazywanej też podwójnym rozczłonkowaniem lub podwójną artykulacją<sup>9</sup> (Polański, 2003). Polega ona na tym, że każda wypowiedź językowa zbudowana jest z ograniczonego zestawu elementów nieposiadających znaczenia (fonemów), które z kolei daje się złożyć w nieograniczony, choć w danym momencie skończony zbiór jednostek obdarzonych znaczeniem (morfemów, wyrazów, grup wyrazowych). Dzięki takiemu strukturalnemu zróżnicowaniu język jest efektywnym narzędziem w przekazywaniu dużej ilości informacji (Hockett, 1968), gdyż przy

---

9 Cecha podwójnej lub dwójakiej artykulacji (fr. *la double articulation*) została zaproponowana przez Martineta (1949, za: Grzegorzczkowska, 2007). Generalnie odpowiada ona temu, co Hockett (1968) nazywał dwoistością, czy dwustopniowością struktury języka (ang. *duality of patterning*), przy czym za pierwszy poziom (pierwsze rozczłonkowanie) Martinet uważał monemy, czyli jednostki posiadające znaczenie (wyrazy, morfemy), z których dają się wyodrębnić jednostki drugiego poziomu (drugiego rozczłonkowania), czyli nieposiadające znaczenia fonemy (Polański, 2003). Hockett (1979) z kolei elementarne jednostki nieposiadające znaczenia nazywał za Hjelmslevem cenemami, natomiast zbudowane z nich morfemy i wyrazy, czyli jednostki znaczące – pleremami. Więcej o podobieństwach i różnicach w ujęciu tej cechy przez Martineta i Hocketta pisze Ladd (2012).

zachowaniu stosunkowo niewielkiego i niezmiennego zestawu jednostek elementarnych pozostaje systemem otwartym na kreowanie nowych jednostek wyższego rzędu. Dotyczy to nie tylko powstawania nowych wyrazów, którymi można określić nieznane wcześniej w danej społeczności przedmioty i zjawiska, ale jeszcze w większym stopniu wiąże się z generowaniem unikatowych, a jednocześnie zrozumiałych dla otoczenia zdań. Można tu więc mówić nie tylko o otwartości tego systemu, ale także o jego produktywności (Hockett, 1968, 1979), czyli możliwości tworzenia nieskończonej liczby zdań ze skończonego (w danym momencie) zbioru wyrazów. Tak rozumiana produktywność języka jest możliwa dzięki jego złożonej strukturze zapewnianej nie tylko przez podwójną artykulację znaku językowego, ale także przez podobną do niej cechę dwuklasowości języka<sup>10</sup>, która została już przywołana w definicji Grzegorzyczykowej (2007). Zasada dwuklasowości języka została sformułowana przez przedstawiciela praskiej szkoły strukturalistycznej, K. Bühlera (1934/2004), który wyróżnił w języku (co najmniej) dwie klasy tworów językowych wchodzących w relacje „część-całość”. Są to wyrazy i zdania. W procesie językowego przedstawiania świata wyrazy odnoszą się do poszczególnych jego elementów, zdania natomiast pozwalają na złożenie tych elementów w spójną reprezentację językową. Zdania nie są zbiorem luźno powiązanych wyrazów, lecz wynikiem ułożenia tych wyrazów w odpowiednim szyku dzięki regułom syntaktyki. Wyrazy tworzą więc pierwszą klasę znaków, drugą zaś klasą są normy łączenia ich w struktury, które reprezentowane są przez zdania (Gołąb i in., 1970). Z tych odrębnych klas tworów językowych możliwe jest zbudowanie nieskończonej liczby wypowiedzi wyrażających dowolną myśl, co decyduje o otwartości i produktywności języka, a także o jego abstrakcyjności. Ta ostatnia cecha języka przejawia się w możliwości odnoszenia wyrażen języka nie tylko do sytuacji w której bezpośrednio znajdują się rozmówcy („tu i teraz”), ale także do przedmiotów i zdarzeń odległych w czasie i przestrzeni. W terminologii Hocketta (1968, 1979) właściwość ta nazywana jest przemieszczeniem czasowo-przestrzennym lub odnośnością na dystans.

Nie tylko struktura znaków językowych i budowanych z nich, dzięki regułom gramatyki wyrażen, wyróżnia język na tle innych sposobów komunikacji, czyniąc go plastycznym narzędziem opisu świata. Twórczy charakter języka wynika również ze związków w jakie znaki języka wchodzą ze światem zewnętrznym (rzeczywistością pozajęzykową). De Saussure (1991) sprzeciwiał się traktowaniu języka jako listy terminów, które odnoszą się bezpośrednio do odpowiadających im elementów rzeczywistości

---

10 Dwuklasowość i podwójne rozczłonkowanie języka bywają ze sobą utożsamiane (Grzegorzyczykowa, 2007), jednak według Polańskiego (2003) zasady te nie są identyczne, a jedynie wykazują pewne podobieństwo (por. Polański, 2004).

(przedmiotów, zjawisk, czynności itd.). Według niego znak języka łączy w sobie pojęcie i obraz akustyczny, które mają naturę psychiczną: pojęcie jest umysłową reprezentacją rzeczywistego obiektu, a obraz akustyczny psychicznym odbiciem (zmysłowym wyobrażeniem) dźwięku mowy (fali akustycznej odbieranej jako mowa). W tym ujęciu znak językowy nie łączy więc bezpośrednio rzeczy i jej nazwy, ale odpowiadające im byty psychiczne: pojęcie i obraz akustyczny, które de Saussure (1991, s. 91) zastępuje określeniami „element znaczący (*signifié*) i element znaczący (*signifiant*)”<sup>11</sup>. Jego zdaniem tak rozumiany znak cechuje relacja dowolności między elementem znaczącym i znaczącym oraz linearny charakter elementu znaczącego. O ile druga z tych cech (zwanych przez de Saussure’a zasadami) jest kontrowersyjna i nie znalazła pełnej akceptacji wśród językoznawców (Izert i Pachocińska, 1998), o tyle pierwsza wpisuje się w szerszą filozoficzną refleksję na temat pochodzenia języka i jego związków z rzeczywistością, współcześnie rozstrzyganą zgodnie ze stanowiskiem de Saussure’a<sup>12</sup>.

Dowolność, a więc konwencjonalność znaków języka nierozzerwalnie łączy się z jego społecznym charakterem. Każdy język naturalny jako narzędzie komunikacji funkcjonuje zawsze w obrębie jakiejś społeczności, której członkowie znają i tworzą reguły jego użycia (nie zawsze w pełni świadomie). Według słów de Saussure’a (1991, s. 100) „aby język mógł zaistnieć, potrzebna jest zbiorowość mówiąca”. Znaki danego języka mają charakter konwencjonalny, gdyż o ich znaczeniu i sposobie użycia decyduje właśnie uzus semiotyczny obowiązujący w społeczności użytkowników tego języka. Inaczej mówiąc, znaczenie słów, czy zwrotów używanych w języku ma charakter umowy i nie jest w żaden inny sposób motywowane (np. podobieństwem brzmieniowym<sup>13</sup>).

Można również powiedzieć, że język jest arbitralny, gdyż między formą a treścią nie istnieje żadna pozajęzykowa relacja. Warto w tym miejscu przypomnieć, że naturalne języki

---

11 Ważnym rozwinięciem idei, że między materialną substancją dźwięku, a rzeczywistym obiektem fizycznym pośredniczy pojęciowa i językowa forma, jest dwuplanowy model języka Hjelmsleva (1979). Ze względu na użyteczność tego modelu, jako punktu odniesienia dla różnych modeli procesu mówienia, zostanie on zaprezentowany w dalszej części tego rozdziału.

12 Spór filozoficzny na temat pochodzenia języka i jego związku z rzeczywistością pozajęzykową znany jest od starożytności, kiedy przyjmował postać kontrowersji *physis-thesis*, czyli pytania, czy nazwy z natury i w sposób poprawny są przypisane oznaczanym przez nie przedmiotom, czy też związek ten ustanawiany jest na mocy konwencji, gdyż sam język należy do dzieł ludzkich (*thesei onta*), a nie boskich (*physei onta*) (Prechtel, 2007). Współcześnie panuje konsensus naukowy na temat społecznego (kulturowego) pochodzenia języka, a spory rozgrywają się wokół tego, jak wyglądała ewolucja języka i na ile jego filogenetyczny rozwój opierał się na już istniejących preadaptacjach, a na ile do zaistnienia typowo ludzkiej mowy potrzebne były specyficzne mutacje genetyczne (por. Żywiczyński i Waciewicz, 2015).

13 Wyjątkiem od tej zasady są onomatopeje, oraz wykrzyknienia, które do pewnego stopnia posiadają naturalne umotywowanie (Grzegorzczkowska, 2007). Wyrazy dźwiękonaśladowcze, w których forma dźwiękowa przypomina dźwięk zjawiska, do którego się odnosi, są przykładem ikonizacji obrazowej. Językoznawcy kognitywni wyróżniają także rodzaje ikonizacji manifestujące się w strukturze gramatycznej zdań (Taylor, 2007).

migowe korzystają ze znaków w dużo większym stopniu umotywowanych relacją podobieństwa między znakiem a zastępowanym przez niego obiektem, zjawiskiem, stanem czy czynnością (cecha ikoniczności). Jest to efekt oparcia się tych języków z konieczności na kanale wizualno-przestrzennym (Dunaj, 2015). Typowy dla języków fonicznych kanał głosowo-słuchowy (akustyczny) wymusza poniekąd linearne nadawanie komunikatów i nie sprzyja całościowemu odzwierciedlaniu cech referowanych zjawisk i przedmiotów. Do pewnego więc stopnia, sformułowane przez de Saussure'a zasady dowolności i linearności są ze sobą ściśle powiązane.

Umowność reguł stosowania znaków języka pozwala na jego ewolucję, gdyż pozostawia system otwarty na opis nowych, nieznanych wcześniej obszarów rzeczywistości. Łączy się to ze zjawiskiem twórczego przesunięcia znaczeń (cecha polisemiczności znaków języka), występującego wtedy gdy do opisu jakiegoś zjawiska lub przedmiotu (najczęściej nowego i nienazwanego w danej społeczności, np. wynalazek) używa się wyrażen już istniejących i mających odmienną referencję (Grzegorzczkova, 2007).

Zmiany systemu języka polegają nie tylko na zmianie znaczeń. Ewolucji, której kierunki wyznacza powszechność stosowania danego elementu, podlegają elementy języka i reguły ich stosowania występujące na wszystkich poziomach organizacji języka: od brzmienia głosek (np. zmiana dominacji, któregoś z allofonów), przez odmianę końcówek fleksyjnych, po znaczenia i budowę złożonych wyrażen idiomatycznych.

Powyższy opis nie wyczerpuje wszystkich wskazywanych przez językoznawców cech języka, a jedynie te, które najczęściej przywoływane są jako definiujące język. Badania poczynione przez strukturalistów w pierwszej połowie XX wieku pozwoliły z różnorodności języków świata wyodrębnić wspólne cechy, które decydują o wyjątkowości ludzkiej mowy. Systemowy charakter języka (złożona struktura, w której elementy na każdym poziomie wchodzą ze sobą w relacje opozycji tworząc większe całości o właściwościach innych niż składowe) oraz konwencjonalny (arbitralny) związek między znakami języka a światem pozajęzykowym decydują o możliwości wyrażenia w języku dowolnych treści, nie tylko oderwanych od aktualnej sytuacji, ale także nie mających żadnego oparcia w rzeczywistości. Grzegorzczkova (2007) nazywa to uniwersalnością języka i zauważa jeszcze jedną ważną jego cechę: za pomocą języka (metajęzyka) można opisać również sam system języka, co warunkuje jego samozwrotność.

## 1.2.2 Język w ujęciu gramatyki transformacyjno-generatywnej

Definiowanie uniwersaliów językowych oraz poszukiwanie ukrytej struktury języka, a także tworzenie metajęzyka dla opisu tych zjawisk stały się jednymi z zadań, jakie postawili sobie językoznawcy w drugiej połowie XX wieku. Rozwój logiki i matematyki przyczynił się do podjęcia prób formalizacji języka naturalnego (Polański, 1971), a na znaczeniu zyskał nurt językoznawstwa określany mianem „formalizmu”<sup>14</sup>. Szczególnie silne wpływy na sposób opisu języka zyskała dominująca odmiana formalizmu tzw. gramatyka transformacyjno-generatywna<sup>15</sup>, która zaproponowała ujęcie języka jako kompetencji umysłu do tworzenia poprawnych wypowiedzi (zdań). Był to przejaw szerszych tendencji w nauce II poł. XX wieku (szczególnie w psychologii) nazywanych „rewolucją poznawczą”, które sprzeciwiały się paradygmatowi behawiorystycznemu i postulowały wyjaśnianie ludzkich zachowań (w tym związanych z posługiwaniem się językiem) w kategoriach działania ukrytych procesów umysłowych. Jedną z najbardziej wpływowych postaci tego ruchu stał się Noam Chomsky, twórca gramatyki transformacyjno-generatywnej. Dokładne zreferowanie tego, w jaki sposób kształtowała się myśl językoznawcza (i jej powiązania z psychologią) w tamtym okresie wykracza poza tematykę niniejszej rozprawy, jednak wyodrębnienie się w językoznawstwie nowego paradygmatu „mentalistyczno-kognitywistycznego” (Puppel, 2006), do którego można zaliczyć teorię gramatyki transformacyjno-generatywnej oraz część jej krytyków (np. przedstawiciele gramatyki kognitywnej), przyniosło nowe spojrzenie na język, którego nie sposób tu pominąć. Przede wszystkim zaś na gruncie teorii transformacyjno-generatywnej wypracowano specyficzną terminologię (np. pojęcia struktury głębokiej i powierzchniowej zdania) oraz specyficzne, modułarne rozwiązanie relacji między językiem a umysłem człowieka, do których odwołują się także modele produkcji mowy przedstawione w rozdziale 2. Z tego względu zarys teorii transformacyjno-generatywnej zostanie nakreślony w tym podrozdziale.

Jak już wspomniano, za czołowego i najbardziej wpływowego przedstawiciela nurtu formalnego uważa się Chomsky’ego (Kurcz, 2005; Polański, 2003). Na przestrzeni lat prezentowane przez niego poglądy na język ewoluowały, jednak pewne założenia nieodmiennie kojarzą się z tym badaczem i opracowaną przez niego teorią gramatyki

---

14 W historii współczesnego językoznawstwa wyróżnia się przeciwstawne sobie kierunki formalizmu i funkcjonalizmu (Kalisz, 2001).

15 Generatywizm powstał w opozycji do amerykańskiej wersji strukturalizmu (tzw. dystrybucjonizmu) jednak przejął część jego aparatu pojęciowego i metod. Zamiast więc zerwania z tradycją strukturalistyczną można mówić raczej o uwypukleniu znaczenia kwestii formalnych. Tym co odróżniało założenia teoretyczne Chomsky’ego od amerykańskiej wersji strukturalizmu było mentalistyczne i uniwersalistyczne myślenie o języku w przeciwieństwie do dominującego przed nim podejścia behawiorystycznego (Polański, 2003; MacNeilage, 2005).

transformacyjno-generatywnej w ujęciu tzw. standardowym<sup>16</sup>. Po pierwsze Chomsky definiuje język jako kompetencję idealnego użytkownika języka do budowania zrozumiałych i poprawnych gramatycznie zdań. Według niego język jest to nieskończony zbiór zdań możliwych do wypowiedzenia dzięki „systemowi procesów generatywnych” (Chomsky, 1982, s. 17). Kluczową właściwością kompetencji językowej jest więc możliwość budowania nieskończonego zbioru zdań przy ograniczonym zbiorze reguł. Reguły te Chomsky (1982) przydziela do trzech – tradycyjnie istniejących w językoznawstwie – działów: semantycznego, fonologicznego i składniowego. Semantyka odpowiedzialna jest za nadawanie wypowiedziom znaczeń (interpretacja znaczeniowa). Fonologia zajmuje się określaniem fonetycznej formy wypowiedzi. Zarówno dział semantyczny, jak i fonologiczny służą jedynie interpretacji zdań, wygenerowanych przez dział syntaktyczny, zapewniając ich odpowiedni odczyt (Mecner, 2007). Według Chomsky’ego w pełni generatywnym, a więc odpowiedzialnym za tworzenie zdań składnikiem kompetencji językowej jest składnia, „czyli gramatyka<sup>17</sup> w wąskim sensie tego słowa” (Kurcz, 2005, s. 101). W skład działu składniowego wchodzi reguły struktur frazowych<sup>18</sup>, reguły transformacyjne oraz słownik<sup>19</sup>. Dysponując hasłami słownika dział składniowy przy użyciu reguł jest w stanie wygenerować dowolne poprawne gramatycznie zdanie, niezależnie od sensowności znaczenia jakie ono niesie<sup>20</sup>. Tym co ogranicza możliwość tworzenia zdań nonsensownych są cechy selekcyjne

---

16 Zdaniem znawcy prac Chomsky’ego, Lyonsa (1998, s. 10-11) „(...) wcześniejsze prace Chomsky’ego – te, które jesteśmy dziś skłonni zaliczać do środkowego, czy »klasycznego« okresu jego twórczości (od połowy lat sześćdziesiątych do początku siedemdziesiątych) – miały największe znaczenie dla dyscyplin nielingwistycznych, dostarczając terminologii, która została przejęta przez ich przedstawicieli i służy im stale w dyskusjach nad kwestiami filozoficznymi i psychologicznymi, na które Chomsky jako pierwszy zwrócił uwagę. Nadto (...) w pewnym sensie nawet dla lingwistów wprowadzenie do dorobku Chomsky’ego powinno się koncentrować na okresie wcześniejszym.”

17 Gramatyce w ujęciu transformacyjno-generatywnym przypisywane są także określenia „uniwersalna” oraz „Umysłowa”. Pojęcie gramatyki uniwersalnej odnosi się do założenia o wspólnych dla wszystkich ludzi mechanizmach generowania zdań (struktura powierzchniowa powstaje na skutek transformacji dokonywanych na strukturze głębokiej). Z kolei sformułowanie „Gramatyka Umysłu” podkreśla założenie Chomsky’ego o zlokalizowaniu kompetencji językowej w autonomicznym module umysłu (każdego użytkownika języka) (Kurcz, 2005; Mecner, 2007).

18 Reguły frazowe dotyczą rozpisywania zdań na frazy (syntagmy), a te następnie na ich kolejne składowe: wyrazy i morfemy. Np. zapis  $S \rightarrow NP + VP$  oznacza, że zdanie symbolizowane przez „S” składa się z frazy rzeczownikowej „NP” oraz frazy czasownikowej „VP”. Reguły frazowe są odpowiedzialne za generowanie struktur głębokich (Izert i Pachocińska, 1998; Polański, 2003).

19 Słownik lub leksykon jest zbiorem leksemów (wyrazów) stosowanych w danym języku. Każde hasło w słowniku zawiera informację o przynależnych mu cechach fonologicznych (macierz fonologicznych cech dystyngtywnych), znaczeniowych i składniowych. Cechy te ograniczają możliwość wyboru danej jednostki leksykalnej ze względu na jej łączliwość z innymi wyrazami. Za klasyfikacje jednostek leksykalnych pod względem ich łączliwości z innymi elementami języka odpowiadają tzw. reguły selekcyjne (Izert i Pachocińska, 1998; Mecner, 2007).

20 Przykładem zdania skonstruowanego poprawnie gramatycznie jednak pozbawionego sensu jest słynne zdanie Chomsky’ego (1982, s. 202) „Zielone bezbarwne idee śpią wściekle” (org. *Colorless green ideas sleep furiously*) (por. Kurcz, 2005).

zawarte w hasłach leksykonu, a nie reguły składniowe, których rolą jest budowanie zdań akceptowalnych gramatycznie (por. Izert i Pachocińska, 1998).

Przy założeniu, że kluczową cechą języka (kompetencji językowej) jest „możliwość tworzenia nieskończonego zbioru wyrażen ze skończonej liczby oddzielnych elementów (Chomsky, 2005, s. 140), modelem tak rozumianego języka jest gramatyka (generatywna), która w pełni sformalizowany i eksplicytny sposób dokonuje strukturalnego opisu dowolnego zdania wytworzonego przez użytkownika języka. Inaczej mówiąc gramatyka generatywna nadaje wypowiedziom ich językową formę dzięki funkcji generowania zdań z określonym opisem strukturalnym, to znaczy przyporządkowuje „każdemu z nieskończonego zbioru zdań opis strukturalny, wskazujący, jak owo zdanie rozumie idealny użytkownik języka” (Chomsky, 1982 s. 17). Nie oznacza to jednak, że gramatyka generatywna jest odpowiedzialna za wytwarzanie wypowiedzi, czy też produkcję mowy. Chomsky wyraźnie odróżnia w tej kwestii kompetencję językową, czyli znajomość języka (niekoniecznie uświadomioną umiejętność posługiwania się regułami języka) od wykonania, czyli praktycznego użycia języka w konkretnych aktach mowy. Podział na kompetencję i wykonanie wyraźnie nawiązuje do rozróżnienia na „la langue” i „la parole”, przy czym w postaci zaproponowanej przez Chomsky’ego przybiera postać kartezyjskiego problemu umysł-ciało (MacNeilage, 2005), a – jedynie sygnalizowany przez de Saussure’a – psychiczny aspekt języka<sup>21</sup>, staje się tutaj kwestią centralną. Język w ujęciu generatywnym nie jest już traktowany jako społeczny wytwór zdolności językowej wszystkich członków społeczności mówiącej danym językiem, ale jako sama kompetencja bytująca w umysłach użytkowników języka. Celem analizy języka („wrodzonej »zdolności językotwórczej« istot ludzkich”) ma być odkrycie reguł, które pozwalają nie tylko opisać strukturę tekstów znanych, ale przede wszystkim przewidzieć konstrukcję dowolnego nowego zdania w danym języku (Chomsky, 1982, s. 52; por. Izert i Pachocińska, 1998).

Aby wyjaśnić w jaki sposób generowane przez gramatykę zdania stają się czytelne pod względem znaczeniowym i fonetycznym, Chomsky (1982) używa terminów struktury głębokiej (*deep structure*) oraz powierzchniowej (*surface structure*), odnoszących się do dwóch różnych poziomów reprezentacji zdania. Według niego, struktura gramatyczna zdania, jaką przedstawia się przy użyciu aparatu analitycznego tradycyjnych gramatyk jest tylko powierzchniową, zewnętrzną formą głębszych, bardziej podstawowych i uniwersalnych struktur gramatycznych, które nie są dostępne bezpośredniej obserwacji (Kurcz, 2005). Tym,

---

21 De Saussure (1961, s. 29) pisał „(...) system gramatyczny istnieje potencjalnie w każdym mózgu lub raczej w mózgach ogółu jednostek, gdyż w żadnej z nich nie jest kompletny; w pełni istnieje on tylko w zbiorowości.”

co stanowi właściwą domenę działania składni jest struktura głęboka. Tworzą ją podstawowe znaczniki frazowe<sup>22</sup> (zbiory generowanych przez składnię rzędków<sup>23</sup>), których ciągi Chomsky (1982) nazywa bazą zdania (por. Mecner, 2007). Inaczej mówiąc, na poziomie struktury głębokiej, w wyniku zastosowania reguł frazowych do haseł słownika powstaje baza zdania, składająca się z tzw. rzędków, które formalnie daje się opisać przy pomocy podstawowych znaczników frazowych. W swojej podstawie oprócz zestawu reguł generujących strukturę głęboką zdania składnia zawiera także podskładnik transformacyjny. Jest on odpowiedzialny za stosowanie operacji formalnych, przekształcających strukturę głęboką w powierzchniową. W uproszczeniu można powiedzieć, że wygenerowana przez reguły frazowe (tworzące podstawowe znaczniki frazowe) baza zdania jest następnie tłumaczona przez reguły transformacyjne (operacje formalne także należące do działu składniowego) na strukturę powierzchniową, czytelną dla działów fonologicznego i semantycznego<sup>24</sup> (Mecner, 2007). „Baza zdania odwzorowywana jest w zdanie przez reguły transformacyjne, które ponadto w trakcie tego procesu automatycznie przyporządkowują zdaniu derywowane znaczniki frazowe (włącznie z końcową strukturą powierzchniową)” (Chomsky, 1982, s. 174).

Analizując różne wypowiedzi w językach naturalnych można znaleźć przypadki, w których dwa zdania o innej budowie powierzchniowej mają taką samą bazę na poziomie struktury głębokiej (np. zdania w stronie czynnej i biernej), a także takie, których struktura powierzchniowa jest identyczna (np. zdania wieloznaczne), ale mogą generować ją dwie zupełnie różne struktury głębokie (Izert i Pachocińska, 1998). Polański (1971, s. 91) podaje z kolei przykład wyrażen „czytanie Janka” oraz „czytanie książki”, które powierzchniowo mają podobną budowę, jednak ich powstanie wiąże się z zastosowaniem zupełnie innych transformacji wobec tej samej struktury głębokiej (zdanie jądrowe brzmiałoby: *Janek czyta książkę*).

---

22 Znacznik frazowy – graficznie przedstawiany najczęściej w postaci drzewa derywacyjnego – stanowi jeden ze sposobów formalnego opisu struktury zdania. Podstawowy (wyjściowy) znacznik frazowy odnosi się do najbardziej abstrakcyjnej (głębokiej) struktury zdania i jest z nią utożsamiany (Izert i Pachocińska, 1998; Polański, 2003).

23 Rządki to szeregi wyrazów, fraz lub zdań, którym odpowiadają symbole, np. NP + V, które oznaczają rządki tworzone przez frazę rzeczownikową oraz czasownik. Rządek może być złożony również z pojedynczego elementu, np. symbol S oznacza zdanie, czyli element nadrzędny, który może być zgodnie z regułami frazowymi rozpisany na składowe. Symbol S daje się na zasadzie rekurencji dodawać do podstawowych znaczników frazowych umożliwiając tworzenie nieskończonego zbioru zdań, ale także – potencjalnie – nieskończenie długiego zdania, np. zdanie składające się z wielu zdań równolegle złożonych oddzielonych spójnikiem „i”: S i S i S itd. (Kurcz, 2005; Polański, 2003).

24 Początkowo model standardowy gramatyki generatywno-transformacyjnej zakładał interpretację semantyczną jeszcze na poziomie struktury głębokiej (tzw. forma logiczna), jednak w wersji rozszerzonej modelu została ona przesunięta na poziom struktury powierzchniowej obok interpretacji fonologicznej (Chomsky, 1982; Kurcz, 2005).

Według Chomsky'ego (1982) reguły transformacyjne należą do uniwersaliów językowych. Podobnie, jak zbiór cech dystynktywnych (uniwersalia materialne) stanowi bazę dla fonologii każdego języka naturalnego, tak zestaw reguł transformacyjnych (uniwersalia formalne) określa dostępne operacje, jakie mogą zostać wykonane na bazie zdania, aby uzyskać strukturę powierzchniową. Odkrywanie tych reguł i weryfikowanie ich obecności w różnych językach świata jest ważnym zadaniem, jakie stawia Chomsky przed językoznawstwem. Píše on „(...) głównym zadaniem teorii lingwistycznej winno być wypracowanie takiego opisu uniwersaliów językowych, który z jednej strony nie mógłby być sfalsyfikowany przez faktyczną, rzeczywistą różnorodność języków, a z drugiej strony byłby wystarczająco bogaty i eksplicytny, by wytłumaczyć zarówno szybkość i jednolitość uczenia się języka, jak też uderzającą złożoność i mnogość gramatyk generatywnych, będących produktem tejże nauki” (Chomsky, 1982, s. 48).

Model języka przedstawiany przez tzw. teorię standardową okazał się rozczarowujący – również dla samego Chomsky'ego – pod względem założeń jakie miał za zadanie spełniać. Po pierwsze badania psycholingwistyczne prowadzone z inspiracji tego modelu nie potwierdziły jego realności psychologicznej (Kurcz, 2005). Ponadto analiza języków innych niż angielski wymagała rozbudowania systemu reguł transformacyjnych, komplikując go bardzo. Stało to w sprzeczności z obserwacją tego, z jaką łatwością i w szybkim tempie dzieci rozwijają biegłość w posługiwaniu się dowolnym językiem. Zauważono, że opanowanie tak wielu reguł w tak krótkim czasie byłoby niewykonalne dla dzieci (Mecner, 2007). Dlatego – zdaniem Chomsky'ego (2005, s. 70) – należało przyjąć, że „podstawowa struktura języka jest zasadniczo jednorodna i nie pochodzi z zewnątrz, lecz od wewnątrz”, a ponadto „(...) należało wykazać, że zróżnicowanie reguł to zjawisko wtórne i można znaleźć bardzo ogólne zasady rządzące wszystkimi regułami”. Zgodnie z tymi przesłankami powstała teoria rządu i wiązania lub – bardziej ogólnie – teoria zasad i parametrów (Mecner, 2007). Zakładała ona, że wszystkie języki bazują na wspólnym systemie uniwersalnych zasad, a tym, co różni odmienne języki między sobą, to inne parametry jakie te języki realizują (Bobrowski, 2009). W myśl tej teorii struktura  $D^{25}$  (czyli głęboka) jest generowana przez system bazowy złożony ze słownika i składni. Natomiast składnik transformacyjny wyprowadzający strukturę S (powierzchniową) ograniczony został do reguły „przesuń  $\alpha$ ” (Mecner, 2007), która wyraża uogólnioną zasadę „przemieszczania czegoś dokądś pod pewnymi warunkami” (Chomsky, 2005, s. 72).

---

25 W teorii zasad i parametrów zamiast terminów „struktura głęboka” oraz „struktura powierzchniowa” używa się odpowiednio określeń „struktura D” oraz „struktura S” (Mecner, 2007).

Zmiany jakim podlegała teoria Chomsky'ego można więc podsumować jako upraszczanie reguł przez ich redukcję i uogólnianie, a także rozszerzanie roli składnika semantycznego (Izert i Pachocińska, 1998). W kolejnych odsłonach teorii zwiększała się także rola słownika (Kurcz, 2005). W ostatniej wersji nazywanej „programem minimalistycznym” (Chomsky, 2007) to właśnie hasła słownika umysłowego odgrywają kluczową rolę w tworzeniu syntagm, gdyż zapisane w nich cechy decydują o możliwości ułożenia ciągów wyrazów tworzących zdania (Bobrowski, 2009). Nie rozróżnia się także struktury głębokiej i powierzchniowej, gdyż zarówno forma logiczna, jak i fonetyczna zawarte są w leksykonie, a ich realizacja w wypowiedzi (wyprowadzanie pary dźwiękowo-znaczeniowej) odbywa się w zamkniętym cyklu transformacyjnym obejmującym całe zdanie (Kurcz, 2005; Mecner; 2007). Inaczej mówiąc moduł językowy pośredniczy między systemem pojęciowo-intencjonalnym (C-I, *conceptual-intentional*), a narządami wytwarzania i odbioru mowy (A-P, *articulatory-perceptual*), generując zdania składające się z wyrazów łączących w sobie formę logiczną (znaczenie) oraz fonetyczną (LF, *logical form*; PF, *phonetic form*). Wybranie ze słownika jednostki leksykalne realizują cechy (np. czas i aspekt czasownika), których przekazanie w optymalny sposób odda myśl użytkownika języka (np. komunikat o wydarzeniu, które już się dokonało) ograniczając jednocześnie łączenie ich w sposób nonsensowny (Kurcz, 2005). Warto zauważyć, że takie ujęcie cyklu transformacyjnego wyraźnie zbliża się do procesualnych modeli produkcji mowy, które opisane zostały w rozdziale 2.

### **1.2.3 Język w ujęciu poznawczym (kognitywnym)**

Równoległe do nurtu generatywistycznego, a z czasem w opozycji do niego rozwijał się nurt funkcjonalny językoznawstwa, podejmujący przede wszystkim problemy badawcze związane z użyciem języka w komunikacji (Kalisz, 2001). Językoznawcy o tej orientacji teoretycznej traktują język jako narzędzie służące porozumiewaniu się poprzez reprezentowanie świata (funkcja reprezentatywna lub referencjalna), wyrażanie cech i stanów osoby mówiącej (funkcja ekspresywna) oraz wywieranie wpływu na odbiorcę komunikatu i skłanianie go do pewnych reakcji (funkcja impresywna lub instrumentalna) (Bühler, 1934/2004; por. Kurcz, 1987).

Pewne aspekty podejścia funkcjonalnego szczególnie wyraźnie kontrastują ze scharakteryzowaną wcześniej gramatyką transformacyjno-generatywną. O ile dla formalistów pokroju Chomsky'ego kluczowy jest precyzyjny opis struktury zdania (*sentence*), które

uznają za podstawową jednostkę analizy (por. Chomsky, 1982), językoznawcy funkcjonalni bardziej skłonni są do badania roli jaką dana struktura pełni w komunikacji. Według Kalisza (2001, s. 15) „podstawą językoznawstwa funkcjonalnego jest badanie języka w kontekście jego użycia przez mówiącego i adresata, celów konwersacyjnych uczestników dyskursu i funkcji języka”. Z tego względu jednostką analizy nie są wyłącznie niepowiązane ze sobą zdania, ale wypowiedzi o szerszym zakresie, stanowiące pewną spójną całość, bo realizujące jakąś intencję komunikacyjną i konstruowane w określonym kontekście, np. będące odpowiedzią na pytanie albo komentarzem do sytuacji (Kurcz, 2005).

Z racji większej uwagi przykładanej do użycia języka, refleksja badawcza językoznawstwa funkcjonalnego obejmuje także pragmatykę, która była pomijana w teorii Chomsky’ego. Składnia, traktowana w podejściu formalnym jako jedyny (lub kluczowy) składnik generujący zdania, w podejściu funkcjonalnym traci na znaczeniu na rzecz semantyki oraz pragmatyki (Kalisz, 2001). Ważne miejsce w badaniach funkcjonalistów zajmuje także fonologia. W przeciwieństwie do postulowanych przez Chomsky’ego abstrakcyjnych uniwersaliów formalnych (transformacji przekształcających hipotetyczną strukturę głęboką w obserwowalną strukturę powierzchniową), językoznawcy funkcjonalni uznają jedynie uniwersalia substancjalne, ustalone typologicznie na podstawie badań rzeczywistego użycia języka w mowie (akustyczne i artykulacyjne właściwości głosek realizowanych w różnych językach, pozwalają na zdefiniowanie uniwersalnego zbioru fonemów) (Kalisz, 2001).

Odmienne stanowiska przedstawiciele nurtów formalnego i funkcjonalnego przyjmują także względem przyswajania języka. Generatywiści zajmują się charakteryzowaniem kompetencji językowej idealnego użytkownika języka, która w dużej mierze jest uznawana przez nich za wrodzoną (stanowisko natywistyczne). Nurt funkcjonalny, kładąc nacisk na komunikacyjną funkcję języka, podkreśla dynamiczny proces jego przyswajania ze względu na rolę jaką pełni to medium w interakcjach społecznych (Kalisz, 2001).

Współcześnie do nurtu funkcjonalnego odwołuje się tak zwane językoznawstwo kognitywne<sup>26</sup>, które rozwinęło się bezpośrednio w opozycji do gramatyki transformacyjno-generatywnej (Kalisz, 2001, Langacker, 2009). Językoznawstwo kognitywne, podobnie jak generatywizm jest podejściem mentalistycznym i podobnie jak on sprzeciwia się behawiorystycznemu ujmowaniu języka, jednak w sposób odmienny odpowiada na pytania postawione pierwotnie przez generatywistów o relacje między językiem a strukturami

26 Podobnie jak w przypadku teorii gramatyki generatywno-transformacyjnej, dokładna charakterystyka językoznawstwa kognitywnego w różnych jego odmianach teoretycznych przekracza ramy niniejszej rozprawy. Z tego względu prezentowane w dalszej części rozdziału założenia tego nurtu reprezentują przede wszystkim teorię gramatyki kognitywnej związanej z badaniami R. Langackera.

umysłu. Według językoznawstwa kognitywnego język nie stanowi odrębnego modułu w umyśle, lecz jest ufundowany na innych zdolnościach poznawczych człowieka. Według Taylora (2007, s. 4) wspólnym dla różnych teorii językoznawstwa kognitywnego „jest przekonanie, że język stanowi integralną część ludzkiego poznania oraz, że każda wnikliwa analiza zjawisk językowych powinna być osadzona w kontekście wiedzy na temat zdolności poznawczych człowieka”. Tym co, łączy językoznawstwo formalne w postaci generatywizmu oraz językoznawstwo kognitywne, jest więc zainteresowanie językiem jako przejawem działania ludzkiego umysłu, co pozwala zaliczyć oba podejścia do paradygmatu „mentalistyczno-kognitywistycznego” (por. Puppel, 2006; Taylor, 2007).

Jednocześnie jednak oba nurty dzieli spojrzenie na relacje, jakie łączą język i poznanie<sup>27</sup>. Dla Chomsky’ego gramatyka jest autonomicznym modułem obliczeniowym umysłu, który generuje poprawne zdania (na bazie operowania abstrakcyjnymi algorytmami), a wszelkie błędy wykonania są efektem „czynnika ludzkiego” leżącego poza kompetencją językową (Tabakowska, 1999, s. 8). Podejście kognitywne uznaje natomiast, że gramatyka nie jest „odrębnym i samowystarczalnym systemem poznawczym, lecz „integralną częścią procesów poznawczych” (Langacker, 2009, s. 18). Nie odróżnia się zatem wiedzy językowej od niejęzykowej, gdyż pozostają one we wzajemnych zależnościach w ramach ogólnej architektury umysłu, a „badanie języka sprowadza się do badania ogólnych procesów poznawczych” (Tabakowska, 1999, s. 12; Kalisz, 2001).

Źródeł wskazanych różnic można doszukiwać się w odmiennych odniesieniach teoretycznych. Chomsky (1982) swoje racjonalistyczne i aprioryczne podejście do języka i umysłu w ogóle wywodził od myśli Kartezjusza. Według A. Wierzbickiej (1999) alternatywą dla takiej „lingwistyki kartezjańskiej” jest „lingwistyka leibnizjańska”, która strukturę umysłu odkrywa raczej poprzez badania semantyczne niż spekulacje na temat abstrakcyjnych reguł składniowych generujących zdania niezależnie od niesionego przez nie znaczenia. O ile więc gramatyka transformacyjno-generatywna poszukuje uniwersaliów formalnych w postaci hipotetycznych operacji transformacyjnych, badania czynione w ramach paradygmatu kognitywnego skłonne są bardziej do definiowania tego, co wspólne wszystkim językom przez pryzmat wspólnych wszystkim ludziom zdolności umysłowych (także percepcyjnych i motorycznych) warunkujących typowo ludzki sposób doświadczania świata (por. Kalisz, 2001). Zdaniem Langackera (2009) codzienne doświadczenie somatyczne przestrzeni, ruchu i siły funduje schematy aktywności mentalnej, które stanowią struktury bazowe dla bardziej złożonych konceptów. Należą do nich między innymi tzw. archetypy pojęciowe (np.

---

27 Problem relacji między językiem i jego użyciem w mowie, a innymi procesami poznawczymi zostanie ponownie postawiony w niniejszej pracy w odniesieniu do kontroli procesu mówienia w rozdziale 3.

rozwijane już w pierwszym roku życia pojęcie „rzeczy” czy „przedmiotu”; por. Taylor, 2007), które przez swoje silne zakorzenienie w doświadczeniu należą do uniwersaliów poznawczych (a tym samym językowych).

Analogicznie, zamiast uniwersaliów w postaci wrodzonych algorytmów transformacji składniowych, Wierzbicka (1999, 2011) poszukuje naturalnego metajęzyka semantycznego NSM (*Natural Semantic Metalanguage*), który oddałby leżące u podstaw każdego języka uniwersalne kategorie pojęciowe. Jest to poszukiwanie swoistego „alfabetu ludzkich myśli”, realizujące „leibnizjańskie” podejście do lingwistyki (Wierzbicka, 1999). W wyniku badań semantycznych nad różnorodnymi grupami języków projekt ten doprowadził do opracowania zestawu 64 podstawowych pojęć uniwersalnych, które miałyby stanowić pierwotne i elementarne (tj. nierozkładalne na prostsze) składniki znaczenia, stanowiące podstawę tworzenia szerszych kategorii (tzw. molekuł, por. Goddard, 2012) i klasyfikowania świata. Należą do nich określenia dotyczące relacji przestrzennych: nad, pod, wewnątrz, ale także pojęcia odnoszące się do ludzi i przedmiotów: coś, ktoś, ty, ja. Według Wierzbickiej (2011) są one obecne w każdym języku, choć mogą należeć do różnych części mowy, a ich wykładnikami bywają morfemy, leksemy lub całe frazemy<sup>28</sup>.

Uwidacznia się tu fundamentalne dla językoznawstwa kognitywnego założenie, że język ma przede wszystkim charakter symboliczny, a „wykładnikami tej samej treści mogą być zjawiska zarówno słownikowe, jak gramatyczne, morfologiczne i składniowe” (Bartmiński, 2011, s. 227). Poszczególne składniki, czy też moduły języka nie tylko nie są traktowane jako osobne i autonomiczne, ale tradycyjnie wyznaczane granice między semantyką i pragmatyką lub składnią i morfologią traktowane są jako płynne i jedynie umowne (Kalisz, 2001; Taylor, 2007). Według Langackera (2009, s. 20) „leksykon i gramatyka tworzą kontinuum składające się ze złożonych struktur symbolicznych”, przez co ich opis można zredukować do połączeń formy i znaczenia. Oznacza to, że każde wyrażenie językowe daje się przedstawić w postaci złożenia (1) struktury semantycznej, (2) struktury fonologicznej oraz (3) relacji symbolizacji między obiema strukturami (Taylor, 2007).

Traktowanie języka jako środka łączącego strukturę semantyczną i fonologiczną w sposób, który pozwala na przywołanie jednej przez drugą (Langacker, 2009) wyraźnie nawiązuje do tradycji strukturalistycznych. Zdaniem Tabakowskiej (1999, s. 6) „teoria języka w wydaniu Langackera opiera się na solidnych strukturalistycznych podstawach i jest konsekwentną kontynuacją klasycznego strukturalizmu wywodzącego się od F. de Saussure’a”. Jednocześnie, jak wykazuje Taylor (2007), gramatyka kognitywna w porównaniu

---

28 Frazem – to utrwalone społecznie połączenie wyrazów o regularnym znaczeniu (np. maszyna do pisania), inaczej nazywane „związkiem łączliwym” (Polański, 2003, s. 184).

do teorii de Saussure'a przyznaje strukturalistom semantycznym, jak i fonologicznym większy poziom wzajemnej autonomii. Zarówno struktury fonologiczne jak i semantyczne mają swoją własną organizację i nie są wzajemnie zdeterminowane<sup>29</sup>. Pogląd taki funkcjonował jednak już wcześniej (tj. przed wyodrębnieniem się językoznawstwa kognitywnego) na gruncie językoznawstwa strukturalnego (por. Hjelmslev, 1979).

Wyczerpujące przedstawienie językoznawstwa kognitywnego – podobnie jak podejścia formalnego – wykracza poza cele niniejszej rozprawy, jednak powyższa krótka charakterystyka pozwala na podsumowanie najważniejszych ustaleń XX-wiecznego językoznawstwa na temat istoty języka.

Po pierwsze, zarówno tradycyjne badania strukturalistyczne, jak i współczesne ujęcie kognitywne traktują język jako system symboliczny, łączący dźwięk ze znaczeniem. Systemowy charakter języka ujawnia się w jego złożonej strukturze, w której większe całości posiadające znaczenie (morfemy, leksemy) są zbudowane z mniejszych (fonemów), które tego znaczenia są pozbawione. Dzięki temu język, jako jedyny naturalny system semiotyczny wykazuje podwójne rozczłonkowanie oraz dwuklasowość. Zapewnia to systemowi języka pewną otwartość na nowe konstrukcje przy względnej stabilności i niezmienności jego podstawowych elementów.

Tradycyjnie język dzieli się na podsystemy lub składniki: fonologiczny, semantyczny oraz składniowy. Językoznawstwo kognitywne stoi na stanowisku, że opis języka można sprowadzić do scharakteryzowania organizacji struktur fonologicznych i semantycznych oraz łączącej je relacji symbolizacji. Językoznawstwo formalne uznaje ten opis za niepełny uznając, że składniki fonologiczny i semantyczny pełnią jedynie rolę interpretacyjną i są podrzędne względem składni. Zgodnie z teorią transformacyjno-generatywną to składnia jest jedynie generatywnym (zdolnym do tworzenia poprawnych zdań) składnikiem języka i jej racjonalne modelowanie pozwala na uzyskanie wyczerpującego modelu języka. Współcześni językoznawcy nie są także zgodni co do natury relacji między językiem i umysłem, choć w przeciwieństwie do klasycznych strukturalistów definiują język bardziej w kategoriach mentalistycznych niż społecznych. Generatywiści skłonni są uważać język za odrębny i autonomiczny moduł umysłu, który jest wrodzoną zdolnością do tworzenia nieskończonego zbioru nowatorskich zdań i nazywany kompetencją językową. Kompetencja ta jest wyraźnie odróżniana od wykonania, w czym dostrzec można wyraźną analogię do klasycznego, strukturalistycznego podziału mowy na język (*la langue*) i mówienie (*la parole*). Językoznawstwo kognitywne nie czyni tak wyrazistych podziałów zarówno między językiem

---

29 Problem wzajemnych relacji między poszczególnymi poziomami organizacji języka powróci jeszcze w kolejnych rozdziałach.

i jego używaniem w mówieniu, jak i wobec języka oraz innych zdolności umysłowych człowieka. Według językoznawstwa kognitywnego, język bazuje na innych zdolnościach poznawczych, a codzienne doświadczenie płynące z działania narządów zmysłów kształtuje to, w jaki sposób język opisuje (obrazuje) to doświadczenie.

Język jest narzędziem komunikacji, a jego intencjonalne używanie jest podstawą tworzenia relacji symbolizacji między strukturą semantyczną i fonologiczną. Relacja ta opiera się na zasadzie arbitralności, co oznacza że nie posiada żadnego umotywowania naturalnego, lecz bazuje na konwencji, którą użytkownicy języka muszą poznać (przyswoić) oraz intencjonalnie stosować w mówieniu<sup>30</sup>.

### 1.3 Mówienie – intencjonalna czynność użycia języka

W tradycyjnym językoznawstwie strukturalnym zjawiska związane z językiem, jako pewnym abstrakcyjnym systemem znaków odróżniane są od zjawisk związanych z samą czynnością mówienia. Podział ten, który w swej istocie – jak zauważa MacNeilage (2005) – ma platońskie korzenie, wywarł ogromny wpływ na XX-wieczne badania nad mową. Wyznacza bowiem podział na językoznawstwo wewnętrzne i zewnętrzne. Uwzględnienie tej historycznej perspektywy jest o tyle istotne, że pozwala zrozumieć, dlaczego te dwa przedmioty, tj. struktura języka i mówienie stanowią odrębne domeny badań. W połowie XX w. dychoomia ta, za sprawą Chomsky’ego (1982) przyjęła postać rozróżnienia pomiędzy kompetencją językową użytkowników języka (*competence*), a wykonaniem (*performance*), czyli rzeczywistym używaniem języka. Choć opozycja kompetencja-wykonanie nie jest tożsama z opozycją język-mówienie, to podobnie wprowadza „(...) linię demarkacyjną dzielącą przedmiot badań językoznawcy od przedmiotu badań innych dyscyplin – psychologii, socjologii, czy etnografii” (Tabakowska, 1999, s. 7). Podział ten, choć jest wystarczająco ostry, aby wyodrębnić dwa niezależne przedmioty badań, jest jednocześnie niemożliwy do pełnego utrzymania w rzeczywistym prowadzeniu badań. Odseparowanie języka od mówienia jest możliwe jedynie na płaszczyźnie teoretycznej analizy, jednak to, co jest rzeczywiście dane w badaniu, to ludzka mowa w swej całości: język używany podczas mówienia (lub pisania). Z tego też względu badacze z innych dyscyplin niż językoznawstwo nierzadko przekraczali wyznaczoną granicę podziału pomiędzy domenami badań i

---

30 Nadawca komunikatu chcąc coś zakomunikować odbiorcy (intencja) musi założyć, że użyte przez niego środki (język) są współdzielone przez odbiorcę na bazie znanej konwencji, dzięki czemu przekazany za ich pomocą komunikat będzie czytelny (zrozumiały). Odbiorca natomiast próbuje zrozumieć nadawcę zakładając, że ma on faktyczny zamiar (intencja) przekazania czegoś i używa w tym celu skonwencjonalizowanych środków (język). Podstawowe reguły konwersacji przedstawił Grice (1980).

wypowiadali się także o języku. W tym miejscu nakreślona zostanie krótka charakterystyka mówienia, jaka wyłoniła się z badań nad tym procesem do czasu powstania pierwszych psycholingwistycznych modeli produkcji mowy, które omawiane są w rozdziale 2.

Najkrótsza definicja języka zaproponowana przez de Saussure'a (1991, s. 100) brzmi „język jest to (...) mowa minus mówienie”. Jeśli więc od scharakteryzowanego listą cech zjawiska mowy „odjąć” to, co już powiedziano o języku, możliwe będzie scharakteryzowanie mówienia (por. lista cech mowy wg Hocketta, 1968; 1979). Jeśli systemowi języka można przypisać cechy: abstrakcyjności, społeczności, potencjalności, ograniczoności<sup>31</sup>, psychiczności, to mówienie charakteryzowałyby się cechami przeciwnymi (Grzegorzczkova, 2007):

1. Konkretność – o ile język jest bytem abstrakcyjnym, istniejącym dzięki jego użytkownikom, ale też w dużej mierze niezależnie od nich (kompetencja językowa), każda wypowiedź jest konkretnym aktem wykonywanym przez daną osobę w określonym miejscu i czasie.
2. Jednostkowość – w przeciwieństwie do języka, którego zasób elementów i reguł ich łączenia jest wspólny dla grupy ludzi (społeczności), mówienie opiera się na idiolekcie (wewnętrznej reprezentacji systemu języka w umyśle danej osoby) i jest wyrazem jednostkowej intencji.
3. Aktualność – oznacza, że mówienie jest właściwym sposobem realizacji potencjału tkwiącego w systemie języka.
4. Nieograniczoność – o ile system języka posiada ograniczony w danym momencie zasób elementów (możliwa zmiana i rozszerzanie zasobu elementów języka dokonuje się w sposób ewolucyjny), mówienie nie ma tak zarysowanych granic.
5. Fizyczność – mówienie jest zawsze zdarzeniem fizycznym, przebiegającym na poziomie behawioralnym (artykulacja) i akustycznym.

Opozycja *langue-parole* przeciwstawia więc abstrakcyjny i społecznie wytworzony system temu, co jest jego indywidualną realizacją. O ile więc język do istnienia wymaga społeczności mówiącej tym językiem, to mówienie jest konkretnym i indywidualnym aktem danego przedstawiciela tej społeczności.

Inna współczesna definicja mówienia wskazuje na jego dwa istotne aspekty: po pierwsze „obejmuje ono procesy psychiczne nadawcy, polegające w zasadzie na wytwarzaniu

---

31 System języka w danym momencie historycznym ograniczony jest do określonego zbioru elementów (zestaw fonemów, leksykon) z których można zbudować nieograniczoną liczbę zdań (cecha produktywności). Elementy te mogą potencjalnie ulec zmianie, co czyni język systemem otwartym, jednak w dalszym ciągu ograniczonym.

pewnych treści pojęciowych, związanych z określonymi przedstawieniami dźwiękowymi”, a po drugie „w skład mówienia wchodzi ruchy aparatu mowy człowieka” (Milewski, 2004, s. 5). Wedle tej definicji mówienie można uznać za czynność mającą swoją wewnętrzną, psychiczną stronę związaną z tworzeniem zrozumiałej dla odbiorcy wypowiedzi na podstawie dostępnych w danym systemie języka elementów i zasad ich łączenia oraz stronę zewnętrzną, fizyczną polegającą na wygenerowaniu przy pomocy ruchów aparatu artykulacyjnego takiego zestawu następujących po sobie dźwięków, który będzie mógł zostać odebrany przez słuchacza i zinterpretowany w tym samym systemie języka.

Zdaniem Maruszewskiego (1970) uznanie mówienia za czynność pociąga za sobą konsekwencję przypisania jej tych samych właściwości, jakie ogólnie cechują inne ludzkie czynności. Według charakterystyki przedstawionej przez Tomaszewskiego (1971, s. 117) „najważniejszą cechą czynności jest jej ukierunkowany przebieg, zmierzający do określonego, wyróżnionego skutku, który nazywamy wynikiem.” W przypadku mówienia (nadawania mowy), podstawowym wynikiem tej czynności jest wygenerowanie poprawnej językowo lub co najmniej „zrozumiałej” dla rozmówcy wypowiedzi. Na tym jednak cel mówienia się nie kończy, gdyż w zależności od intencji komunikacyjnej nadawcy może on oczekiwać skutków swojej czynności mownej w postaci zmiany zachowania otoczenia, które odebrało jego wypowiedź. Zwrócił na to uwagę Austin (1962/1993) wykazując, że podczas mówienia czynnościowy charakter ma nie tylko wykonywanie ruchów przez mięśnie aparatu mowy, ale także wyrażanie pewnej intencji oraz osiąganie związanych z tą intencją rezultatów, co pozwala traktować mówienie w kategoriach działania. W związku z tym wprowadził on rozróżnienie tych trzech aspektów mówienia na czynności<sup>32</sup> lokucyjne, illokucyjne i perlokucyjne. Czynności lokucyjne związane są bezpośrednio z wypowiadaniem dźwięków (czynność fonetyczna), które tworzą słowa układane dzięki regułom składni (czynność fatyczna) i posiadające swoje odniesienie do rzeczywistości pozajęzykowej (czynność rematyczna). Czynności lokucyjne (łącznie) odpowiadają więc za treść wyrażanego sądu (Kurcz, 1987). Jest to najbardziej podstawowy sens, w jakim mówienie traktowane jest przez Austina (1962/1993) w kategoriach czynnościowych. Jednocześnie „mówiąc coś” wyraża się pewną intencją, tj. podejmuje pewne działanie, dla którego czynność mówienia (lokucja) jest tylko środkiem realizacji (tę samą intencję można by wyrazić na piśmie, w formie gestu, czy przy pomocy innego konwencjonalnego aktu komunikacji). Ten aspekt

32 Tłumacz prac Austina na j. polski termin „speech act” przekłada jako „czynność mowy”. Dużo bardziej rozpowszechnione jest jednak tłumaczenie „akt mowy” (por. Kucz, 1987; 2005). W związku z tym, że termin „akt mowy” bywa z kolei utożsamiany głównie z jego aspektem illokucyjnym („akt illokucyjny”) (por. Searle, 1980) zdecydowano się pozostać przy oryginalnym tłumaczeniu, podkreślającym czynnościowy charakter mówienia, tym bardziej, że w dalszej części pracy proces mówienia analizowany będzie głównie w jego aspekcie lokucyjnym.

mówienia nazywa Austin czynnością illokucyjną. Wykonanie czynności illokucyjnej odróżnia się od wykonania czynności lokucyjnej w ten sposób, że jest to „wykonanie czynności kryjące się w powiedzeniu czegoś, w przeciwieństwie do wykonania czynności powiedzenia czegoś” (Austin, 1962/1993, s. 645). Poza tym „poprzez mówienie” osiąga się jeszcze jakieś zewnętrzne skutki, czyli wykonuje czynność perlokucyjną, przy czym konieczne dla jej wystąpienia jest odebranie wypowiedzi przez słuchacza. Czynność perlokucyjna może wprost odnosić się do intencji osoby mówiącej, ale może też być z nią niepowiązana, gdyż „to, co stanowi perlokucyjny cel jednej illokucji, może być następstwem innej” (Austin, 1962/1993, s. 663), a także osiągnięty dzięki mówieniu rezultat może być zupełnie niezamierzony (np. onieśmienie kogoś podczas wyrażania pochwały).

Oprócz specyficznego ukierunkowania, za możliwość zrealizowania wyniku, do którego zmierza dana czynność, odpowiada jej wewnętrzna struktura (Tomaszewski, 1970). Zorganizowany charakter czynności pozwala uznać je za pewne odrębne od otaczających procesów złożone całości, których części składowe wchodzi z sobą w relacje określonego typu (Maruszewski, 1970). Struktura mówienia ma charakter linearny, co wynika po części z korzystania z kanału głosowo-słuchowego i wymaga wygenerowania dźwięków mowy w określonym porządku. Nawet niewielka zmiana w szyku głosek, czy wyrazów może diametralnie zmienić znaczenie wypowiedzi i stać się zakłóceniem w przebiegu czynności. Ponadto mówienie zazwyczaj odbywa się w jakiejś określonej sytuacji komunikacyjnej, w której – jeśli zakładana jest wymiana komunikatów – rozmówcy zachowują sekwencję naprzemienności mówienia (i słuchania), wyznaczającą strukturę konwersacji. Struktura mówienia określana jest także przez pewne konwencjonalne zasady, wykraczające poza reguły użycia języka, a odnoszące się do zjawisk społecznych, np. specyficzne zwroty wynikające z hierarchii społecznej, słowa zakłóć lub wypowiedzi towarzyszące rytuałom itp.

Antycypowany skutek czynności, na który jest ona ukierunkowana jest także punktem odniesienia dla oceny jej przebiegu i skuteczności. Nie tylko więc wynik ukierunkowuje czynność i organizuje jej przebieg, ale też wyznacza standard, wedle którego możliwe jest porównanie rozbieżności między stanem oczekiwanym a obserwowanym (Łukaszewski, 2002). W przypadku wystąpienia utrudnień (ograniczających lub zupełnie uniemożliwiających osiągnięcie określonego stanu końcowego) możliwe jest podjęcie działań wyrównawczych (np. reperacyjnych), przywracających ukierunkowanie na pierwotny wynik (Tomaszewski, 1971). Ten sam mechanizm korekcyjny można przypisać mówieniu. Jeśli wynik czynności nadawania mowy nie zostaje uzyskany, mogą nastąpić zmiany na różnych poziomach jej organizacji, począwszy od powtórzenia tego samego komunikatu, ale z większą

intensywnością głosu, po dokonania istotnych przekształceń w strukturze samej czynności (Maruszewski, 1970). Podobnie Austin (1962/1993) zauważa, że każdy z aspektów aktu mowy (lokucyjny, illokucyjny, perlokucyjny) może wiązać się z niepowodzeniem i niezrealizowaniem pierwotnego zamiaru. Dlatego podczas mówienia nadawca dokonuje porównań tego co mówi, z tym co zamierzał powiedzieć oraz z tym, co przez mówienie zamierzał osiągnąć. Kluczowe znaczenie w tym procesie ma intencja nadawcy, gdyż to ona wyznacza cele czynności mówienia zarówno związane z bezpośrednim efektem, czyli tekstem wypowiedzi (co mówię?), znaczeniem tej wypowiedzi (co chcę przez to powiedzieć? jakie znaczenie chcę przekazać?) oraz dalszymi konsekwencjami w postaci reakcji odbiorców (jakiej zmiany w funkcjonowaniu otoczenia oczekuję?).

Szczegółową analizę relacji między pojęciami intencji i jej realizacji, a procesem mówienia przeprowadził Frydrychowicz (1999). Wynika z niej m. in. wniosek, że realizacja intencji pełni nadrzędną i integrującą rolę względem mówienia, gdyż reprezentuje ona cele tej czynności, przez co pozwala ją ukierunkować i zweryfikować zgodność realizacji (wykonania) z leżącym u jej podłoża planem. Jednocześnie to właśnie plan, a nie sama intencja może być uznany za bezpośrednią reprezentację celu mówienia i procesu dążenia do tego celu (Frydrychowicz, 1999). Dzięki temu rozróżnieniu między intencją, a planem można zauważyć, że podczas porównania zgodności i rozbieżności występujących między stanem oczekiwanym a obserwowanym wynikiem czynności mówienia, plan procesu mówienia stanowi właściwy standard dla tych porównań. Rozważając zatem kluczowy dla niniejszej pracy problem kontroli procesu mówienia w kontekście przejęzyczeń i innych błędów mowy, autokorekcje podejmowane przez nadawcę wynikają właśnie z wykrycia niezgodności między planem i jego wykonaniem. Oczywiście działania naprawcze mogą być podejmowane także w wyniku dostrzeżenia trudności w zrozumieniu intencji nadawcy przez interlokutora (utrudnienie w aspekcie illokucyjnym) oraz braku oczekiwanych efektów w otoczeniu (utrudnienie w aspekcie perlokucyjnym). Jednakże reperacje spontanicznych błędów mowy następują w tak krótkich odstępach po wykryciu zakłócenia, że ich źródeł należy upatrywać właśnie w procesach monitorowania i porównywania wyniku, czyli tekstu wypowiedzi z jej wewnętrznym planem. W wypadku wystąpienia i wykrycia niezgodności jest ona niezwłocznie korygowana (przywracanie ukierunkowania na wynik) bez oczekiwania na jej krótko i długofalowe skutki. Dlatego też możliwe jest rozpatrywanie kontroli procesu mówienia już w odniesieniu do czynności lokucyjnych, niejako w oderwaniu od czynności illokucyjnych i perlokucyjnych (zob. rozdział 3.) Jednocześnie wykrywana przez nadawcę niezgodność wykonania względem planu, zawsze jest niezgodnością w stosunku do intencji,

dlatego też to intencja jako „integrator” procesu mówienia (Frydrychowicz, 1999) jest traktowana, jako ostateczna instancja weryfikująca stopień osiągnięcia wyniku i ukierunkowująca ewentualne zmiany w czynności mówienia (np. poprzez zmianę planu).

Problem „wewnętrznej”, psychicznej strony mówienia obejmującej m. in. rozważane powyżej procesy (opracowanie intencji, przygotowanie planu wypowiedzi) i jej przełożenia na stronę „zewnętrzną” (fizyczną) zostanie przedstawiony w kolejnym rozdziale w odniesieniu do poznawczych modeli produkcji mowy. Modele te oprócz opisu składających się na proces mówienia poziomów przetwarzania informacji (wyrażającej treściowy aspekt intencji komunikacyjnej), postulują także mechanizmy regulujące przebieg generowania mowy w sytuacji wystąpienia zakłóceń (np. błędów). Zagadnienie to szerzej omówione zostanie w rozdziale 3.

## 2 Psycholingwistyczne modele procesu mówienia

Według Searle'a (1980), podstawową jednostką komunikacji językowej nie jest symbol, słowo ani zdanie, lecz ich wyprodukowanie w konkretnym wykonaniu aktu mowy. Produkcja ta obejmuje nie tylko wytworzenie określonego zestawu dźwięków, ale także ich formalne zorganizowanie, aby stały się czytelne dla odbiorcy znakami języka, reprezentującymi rzeczywistość pozajęzykową i realizującymi intencję nadawcy. O ile więc w systemie języka tkwi potencjał komunikacyjny niespotykany w innych naturalnych sposobach komunikacji, to jego realizacja w mowie (aktualizacja potencjału) musi uwzględniać także jego złożoność na wszystkich poziomach organizacji (wszystkie podsystemy języka).

Dla zilustrowania tego problemu przydatny będzie schemat znaku językowego, zaproponowany przez kontynuatora myśli de Saussure'a i twórcę glossematyki Hjelmsleva (1979)<sup>33</sup>. Według tego językoznawcy znak językowy można podzielić na dwie płaszczyzny: plan treści i plan wyrażania (rysunek 1). Każda z płaszczyzn składa się z kolei z charakterystycznej dla siebie substancji i formy. Substancją treści jest całość zjawisk i przedmiotów poznawalnych przez ludzki umysł, które dają się wyrazić dzięki temu, że można nadać im odpowiednią formę. Forma treści jest natomiast tym, co organizuje (kategoryzuje) kontinuum poznawalnego świata w postaci niepodzielnych jednostek (figur). O ile więc substancja treści odnosi się do realnego świata, o tyle forma treści jest jego pojęciowym obrazem, do którego należą na przykład opozycje pomiędzy tym co wysokie i niskie (Fisiak, 1985; Eco, 2002). Zależność między formą i substancją planu treści obrazuje Hjelmslev (1979) odwołując się do przykładu kolorów: kontinuum widma barwnego (substancja) jest odmiennie kategoryzowane przez formę treści poszczególnych języków.

Podobna analogia zachodzi w planie wyrażania. Substancją wyrażania są tu zewnętrzne, dostrzegalne zmysłowo efekty działania formy, która z pewnego kontinuum (fala dźwiękowa) możliwych dźwięków wybiera te, które w danym języku stanowią narzędzie przekazania treści, jednocześnie wykluczając te, które – choć możliwe do wyartykułowania – w skład dźwięków danego języka nie wchodzi (Eco, 2002).

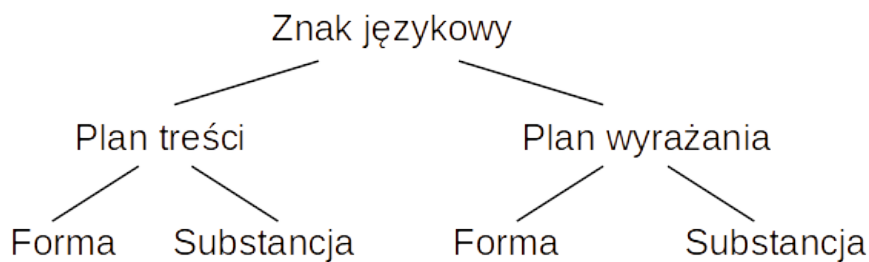
Według Hjelmsleva (1979) język łączy w sobie obie płaszczyzny: treści myśli nadawcy dają się wyrazić za pomocą dźwięków mowy, a usłyszane dźwięki mowy niosą znaczenie formułujące jakąś myśl w umyśle odbiorcy. Jednak istniejąca pomiędzy formą

---

33 Według Kalisza (2001) Hjelmsleva można zaliczyć do przedstawicieli językoznawstwa funkcjonalnego.

## Rysunek 1

### Znak językowy według Hjelmsleva



Źródło: Frydrychowicz (1999)

wyrażania i formą treści współzależność nie ma charakteru bezwzględnie symetrycznego, a jednostki form obu planów nie są do siebie w sposób niezmienny przypisane (Fisiak, 1975). Uznanie względnej autonomii między planami treści i wyrażania<sup>34</sup>, bliskie jest współczesnemu, kognitywnemu ujęciu znaku, jako relacji symbolizacji między znaczeniem, a dźwiękiem (Taylor, 2007).

Dla mówienia, które ma aktualizować potencjał języka, brak precyzyjnej odpowiedniości między planem treści (organizacja semantyczna) i planem wyrażania (organizacja fonologiczna) oznacza konieczność ciągłego ustalania tych relacji w każdym akcie mowy. W ten sposób mówienie można rozumieć nie tylko jako czynność generowania dźwięków mowy (substancji), ale także każdorazowe poszukiwanie właściwej formy treści i jej adekwatnego tłumaczenia na formę wyrażania. Problem ten w języku potocznym występuje jako „poszukiwanie brakujących słów dla wyrażenia myśli”. Należy bowiem pamiętać, że mówienie – choć rozpatrywane jako realizacja języka – nie służy językowi, lecz komunikacji. Istotna jest zatem nie tylko aktualizacja potencjału języka, ale przede wszystkim realizacja intencji komunikacyjnej (por. Frydrychowicz, 1999), która stanowi podstawę relacji symbolizacji między znaczeniem i dźwiękiem. Według Taylora (2007, s. 52) „istota symbolu polega na tym, że używa się go z intencją symbolizacji pewnego pojęcia”. Tym samym mówienie traktowane może być jako czynność złożona, wykraczająca poza domenę samej mowy i obejmująca również operacje poznawcze dokonywane przed wygenerowaniem ciągu dźwięków („wewnętrzna” strona mówienia).

34 Problem relacji między planem treści, a planem wyrażania będzie rozważany ponownie w rozdziale 4. w kontekście ontogenetycznych źródeł procesu mówienia (por. Frydrychowicz, 1988; Levelt, 1998).

Zgodnie z powyższymi uwagami, w niniejszym rozdziale omówione zostaną propozycje teoretyczne ujmujące mówienie w sposób procesualny.

## 2.1 Płaszczyzny mowy według Wygotskiego

Na początku zaprezentowany zostanie schemat płaszczyzn mowy, opracowany przez L. S. Wygotskiego (1934/1989). Jego współczesna wartość ma charakter głównie historyczny. Powstał on na długo zanim rozwinęła się współczesna psycholingwistyka i ze względu na uwarunkowania polityczne nie wpłynął bezpośrednio na kształt najważniejszych poznawczych modeli produkcji mowy. Jednocześnie, zachowując odrębną terminologię oraz sposób konceptualizacji procesu mówienia, wykazuje pewne punkty styczne z powstałymi później modelami psycholingwistycznymi. Szczególnie interesujące jest zaproponowane przez Wygotskiego pojęcie mowy wewnętrznej, której badanie może być kluczowe dla zrozumienia zarówno werbalnej strony myślenia, jak i przedartykulacyjnych mechanizmów kontroli mówienia (por. Geva i Fernyhough, 2019). Ponadto przedstawienie tego modelu jest także istotne ze względu na rozważania teoretyczne na temat rozwoju mowy i języka w dzieciństwie, podejmowane w rozdziale 4.

Wygotski wychodzi od rozróżnienia dwóch zjawisk: mowy zewnętrznej i mowy wewnętrznej. Pierwszą różnicą dotyczącą tych dwóch typów mowy, są ich behawioralne wskaźniki. O ile mowa zewnętrzna jest w pełni jawnym zachowaniem, a jej obserwowalnym przejawem są ruchy narządów artykulacyjnych wywołujące dźwięk, o tyle mowa wewnętrzna pozostaje zachowaniem prywatnym, ukrytym dla zewnętrznego obserwatora.

Wygotski uważał jednak, że te dwie odmienne postaci mowy dzielą dużo głębsze różnice: „jak mowa wewnętrzna nie jest mową minus dźwięk, tak mowa zewnętrzna nie jest mową wewnętrzną plus dźwięk” (Wygotski, 1934/1989, s. 396). Według niego oba rodzaje mowy różnią się istotnie zarówno pod względem funkcji, jak i struktury. Po pierwsze mowa zewnętrzna służy porozumiewaniu się ludzi i można utożsamić ją z pojęciem mowy definiowanym w pierwszym rozdziale niniejszego opracowania. Mowa wewnętrzna natomiast jest według Wygotskiego narzędziem poznania, specyficznym rodzajem myślenia określanym przez niego jako „myślenie werbalne”. O ile więc funkcją mowy zewnętrznej jest wykorzystanie języka do komunikacji, o tyle mowa wewnętrzna ma swoją funkcję w wykorzystaniu języka jako narzędzia myślenia. Oprócz odmiennych funkcji mowa wewnętrzna i zewnętrzna mają także różną strukturę, dlatego mówienie (czyli „przejście od mowy wewnętrznej do zewnętrznej”) nie jest jedynie „zwykłym dołączeniem fonii do mowy

wewnętrznej, nie jest zwykłą wokalizacją mowy wewnętrznej, lecz (...) jest złożoną, dynamiczną transformacją, jest przekształceniem mowy predykatywnej i idiomatycznej w mowę syntaktycznie rozczłonkowaną i zrozumiałą dla innych” (Wygotski, 1934/1989, s. 396).

Wygotski, podobnie jak współcześni badacze procesu mówienia, postuluje istnienie kilku poziomów lub etapów przechodzenia od mowy wewnętrznej do zewnętrznej, które nazywa płaszczyznami: „myślenie werbalne rysuje się nam jako złożona dynamiczna całość, w której stosunek między myślą a słowem określić się daje jako ruch poprzez wiele wewnętrznych płaszczyzn, jako przejście z jednej płaszczyzny na drugą” (Wygotski, 1934/1989 s. 404).

Zanim jakaś myśl zostanie wyrażona w konkretnych słowach, musi – według Wygotskiego – zaistnieć motyw, czyli powód dla którego dana myśl ma zostać wypowiedziana. Motywy są tłem afektywno-wolicjonalnym myśli, a u ich źródeł leżą potrzeby, zainteresowania, pobudki, a także afekty i emocje. Motyw, jako pierwsza płaszczyzna mowy, rozpoczyna cały proces mówienia, poprzez generowanie i kierowanie biegiem myśli. Płaszczyzna myśli, odpowiada niewerbalnej sferze pojęciowej. Wygotski twierdził, że „myśl zawsze jest czymś całościowym, czymś znaczenie większym zakresowo i przestrzennie niż pojedyncze słowo” (Wygotski, 1934/1989 s. 400). Na tej płaszczyźnie kształtowany jest temat wypowiedzi, który potencjalnie może zostać przekazany za pomocą różnych słów, tak samo jak różne myśli mogą zostać ujęte w tych samych słowach. Oznacza to brak bezpośredniego przełożenia myśli na konkretne słowa. Według Wygotskiego myśl „(...) nie wyraża się w słowie, lecz spełnia się w nim” (Wygotski, 1934/1989 s. 401). Oznacza to, że dopiero na poziomie słów myśli uzyskują relewantną formę, która nie była wcześniej zdeterminowana. Inaczej mówiąc: myśli nie tylko nie są tożsame ze słowami, ale także nie odpowiadają ich bezpośredniemu znaczeniu, które zajmuje płaszczyznę pośrednią między myślą a słowem. Zanim więc proces mówienia osiągnie płaszczyznę słów, przechodzi przez płaszczyznę znaczeń. Zdaniem Wygotskiego 1934/1989, s. 401) „myśl nie tylko jest zewnętrznie upośredniona znakami, lecz również wewnętrznie jest upośredniona znaczeniami”. Na tej płaszczyźnie znaczeń przebiega właściwe myślenie werbalne, czyli mowa wewnętrzna. Jej jednostkami strukturalnymi są znaczenia słów. Jest ona zarazem „(...) mową, a więc myślą związaną ze słowem”, oraz „(...) w znacznej mierze myśleniem czystymi znaczeniami, (...) momentem dynamicznym, nietrwałym i płynnym” (Wygotski, 1934/1989, s. 397). Taka mowa, służąca przede wszystkim organizacji własnego działania, ma strukturę predykatywną, ze składnią znaczeń dominującą nad gramatyką języka. Jako płaszczyzna pośrednia, wiąże ze sobą w dynamicznym procesie myśli i słowa.

Wygotski wyróżnia jeszcze jedną płaszczyznę pośrednią, zajmującą miejsce pomiędzy słowem wewnętrznym („czystym znaczeniem”) a słowem zewnętrznym (artykułowanym). Jest to płaszczyzna znaczenia słowa zewnętrznego, na której zachodzi transformacja składni znaczeń w stronę składni słów, co w konsekwencji powoduje przystosowanie wypowiedzi nadawcy do potrzeb i możliwości odbiorczych słuchacza. Strukturalnie jest to już mowa zewnętrzna, pozbawiona jednak swojej akustycznej realizacji.

Ostatnim etapem procesu mówienia jest poziom słów. Według Wygotskiego słowo jest jednostką płaszczyzny mowy zewnętrznej. Jako materialny wytwór procesów zachodzących na wcześniejszych płaszczyznach przejawia się w formie fonicznej.

Warto zwrócić uwagę, że Wygotski (1934/1989) nie zakładał jednokierunkowego przepływu między płaszczyznami, lecz wskazywał na swobodne przechodzenia informacji między poziomami, z możliwością przerwania tego procesu w dowolnym momencie, bez konieczności uzewnętrznienia jego rezultatów. Jest to o tyle istotne, że charakter relacji między płaszczyznami, czy poziomami produkcji mowy będzie jednym z najostrejszych sporów między twórcami poznawczych modeli, które omówione zostaną w dalszej części niniejszego rozdziału (por. Harley, 2001; Levelt, 1993; 1999; Lind i Hartsuiker, 2020; Nozari, 2020; Roelofs, 2020a; Vigliocco i Hartsuiker, 2005).

## 2.2 Poznawcze modele produkcji mowy

We współczesnej psycholingwistyce badania nad przetwarzaniem języka (*language processing*) oprócz zagadnień związanych z jego nabywaniem (*language acquisition*) obejmują przede wszystkim kwestie generowania wypowiedzi słownej podczas mówienia (ang. *speech production*, *language production*)<sup>35</sup> oraz jej odbioru i rozumienia (np. ang. *speech perception*, *speech comprehension*, *word recognition*) (por. Bock, 1996; Dell, 1986). W związku z tym, że procesy te w dużej mierze przebiegają w sposób nieobserwowalny (wewnętrznie) zaproponowano szereg modeli opisujących przetwarzanie informacji językowej w architekturze systemu poznawczego. Choć do pewnego stopnia procesy te przebiegają analogicznie, nie można z pewnością skonstatować, że mówienie jest odwrotnością słuchania i rozumienia (por. Levelt, 2001a).

---

35 Używane w anglojęzycznej literaturze wyrażenia „speech production” i „language production”, są w niniejszej pracy tłumaczone jako „produkcja mowy” lub „generowanie wypowiedzi” i stosowane wymiennie z kluczowym terminem „proces mówienia”, które ma ugruntowaną pozycję w polskim piśmiennictwie psycholingwistycznym (por. Frydrychowicz, 1999). Na oznaczenie procesów cząstkowych, zachodzących przed artykulacją używane również będą terminy „programowanie (mowy)” i „kodowanie (językowe)”.

W niniejszym podrozdziale przedstawione zostaną psycholingwistyczne modele produkcji mowy (*speech production*), które na podstawie danych obserwacyjnych (spontaniczne wypowiedzi oraz pojawiające się w nich błędy) i eksperymentalnych próbują wyjaśnić przebieg procesu mówienia. Modele te wywodzą się z klasycznego w psychologii poznawczej podejścia mikrogenetycznego (Levelt, 1998), które umysłową aktywność człowieka przedstawia w postaci ukrytego procesu prowadzącego do zewnętrznych efektów behawioralnych. W przypadku procesu mówienia, modele produkcji mowy postulują szereg poziomów lub etapów, na których wypowiedź (pojedynczy wyraz lub całe zdanie) jest generowana aż do jej wyartykułowania. Levelt (1998, s. 169) nazywa to „umysłową mikrogeneracją słów” (*mental microgenesis of words*).

Ważnym zadaniem, jakie stoi przed teorią opisującą przebieg procesu mówienia jest wyjaśnienie w jaki sposób substancja treści, będąca wyrazem myśli, zostaje ubrana w językową formę i wyrażona w substancji planu wyrażania (fali dźwiękowej). Pytanie, jakie należy sobie postawić w pierwszym rzędzie dotyczy tego, czy elementy strukturalne języka wyróżniane przez językoznawców są realnymi jednostkami programowania mowy oraz w jaki sposób w trakcie mówienia uwzględniane są reguły łączenia tych elementów w ramach systemu semantycznego, syntaktycznego i fonologicznego. Inaczej mówiąc, teoria produkcji mowy powinna określić na ile strukturalny, statyczny opis języka odzwierciedla rzeczywiste procesy konstruowania wypowiedzi językowej, zachodzące w dynamicznej sytuacji mówienia. Według D. B. Fry’ a (1969) szczegółowa analiza przejęzyczeń, dostarcza dowodów na to, że wiedza użytkowników języka o języku jest zbieżna z jego naukowym, językoznawczym opisem, a reguły łączenia elementów systemu (wyrazów, morfemów, fonemów) są regułami które organizują przebieg kodowania wiadomości podczas mówienia. Autor ten zwrócił uwagę na przejęzyczenia, gdyż nie są one przejawem braku określonej wiedzy lub kompetencji językowej, ale chwilowym zakłóceniem działania ukrytych mechanizmów programowania mowy. Mechanizmy te do pewnego stopnia można zrekonstruować analizując regularności ujawniane w dużych zbiorach spontanicznych przejęzyczeń. Idea, że spontaniczne pomyłki w mowie mogą być „oknem na świat” wewnętrznych procesów odpowiedzialnych za mówienie pochodzi jeszcze z końca XIX w., kiedy to Rudolf Meringer (psychiatra) i Carl Mayer (językoznawca) opublikowali pierwsze systematyczne opracowanie dotyczące spontanicznie powstających przejęzyczeń (Cutler, 2006). Zdaniem Fry’ a (1969) dane dostarczane przez analizę korpusów przejęzyczeń pozwalają na wyróżnienie pięciu poziomów kodowania wypowiedzi językowej, które odpowiedzialne są za transformację myśli w ciąg artykułowanych dźwięków mowy. Na

każdym z tych poziomów komunikat reprezentowany jest w specyficzny sposób, odpowiadający jednostkom i regułom ich łączenia związanym z odmiennymi poziomami systemu języka.

Według Fry'a (1969) należy wyróżnić poziomy:

1. kodowania semantycznego – związanego z myśleniem o tym co powiedzieć, ale jeszcze nie w postaci konkretnych słów (pojęciowy charakter reprezentacji);
2. kodowania leksykalnego – na którym dokonywany jest wybór wyrazów, z których składać się będzie wypowiedź; wyrazy reprezentowane są tu w postaci podstawowych jednostek leksykalnych (rdzeni);
3. kodowania morfologicznego – na tym poziomie do morfemów leksykalnych (rdzennych) z poprzedniego poziomu dołączone zostają morfemy składniowe i słowotwórcze, dobierane ze względu na reguły gramatycznej odmiany i tworzenia wyrazów;
4. kodowania fonologicznego – na którym wypowiedź jest reprezentowana w postaci ciągu fonemów odpowiadających układowi morfemów z poprzedniego poziomu.
5. kontroli motorycznej – związany z przekazywaniem instrukcji do mięśni odpowiedzialnych za oddychanie, fonację i artykulację, aby zrealizowały program wyznaczony przez kodowanie fonologiczne.

Wyróżnienie poziomów kodowania wypowiedzi, pozwoliło potraktować mówienie tak, jak inne złożone i uporządkowane zachowania, które przed swoją zewnętrzną manifestacją reprezentowane są w postaci wewnętrznych planów lub programów (por. Dell, 1986, Miller, Galanter, Pribram, 1980). Ponadto, zaproponowane poziomy kodowania wypowiedzi, Fry (1969) powiązał z jednostkami systemu językowego, zgodnymi ze strukturalnym opisem języka. W ten sposób statyczny model języka, traktowany jako wiedza, czy też kompetencja językowa, udało się wpisać w proces wykorzystania tej kompetencji w generowaniu wypowiedzi. Odwołując się do wspomnianej w pierwszym rozdziale terminologii glossematycznej Hjelmsleva (1979), można powiedzieć, że dzięki wielopoziomowym operacjom kodowania, realizowanym podczas mówienia plan treści zostaje zrealizowany przez plan wyrażania (por. rysunek 1)

Propozycja Fry'a (1969), choć bardzo ogólnikowa, stała się punktem odniesienia dla powstających w kolejnych latach modeli produkcji mowy. Samo zwrócenie uwagi na to, że błędy mowy mogą dostarczyć informacji o organizacji i przetwarzaniu języka w umyśle, stało w opozycji do generatywistycznej idei opisu kompetencji idealnego, a więc bezbłędnego użytkownika języka (Baars, 1992; por. Chomsky, 1985). Inni badacze zajmujący się tym

zagadnieniem, podobnie jak Fry, zaczęli korzystać z danych dostarczanych przez analizę spontanicznie pojawiających się przejęzyczeń, a także projektowali eksperymenty (Levelt, 1999a), mające zweryfikować adekwatność tego opisu również dla procesu mówienia niezaburzonego pomyłkami. W ten sposób powstało wiele modeli produkcji mowy, które w większym stopniu precyzowały kwestie zarysowane jedynie w ogólny sposób przez Fry'a (1969), np. problem relacji między poziomami kodowania (jednokierunkowy przekaz czy interakcja; przetwarzanie sekwencyjne czy kaskadowe). Zaprezentowanie wszystkich tych modeli z dokładnym opisem wyróżniających je założeń wymagałoby osobnego opracowania, dlatego w dalszej części rozdziału przedstawione zostaną modele, które najczęściej dyskutowane są w literaturze psycholingwistycznej (por. Frydrychowicz, 1999; Levelt, 1999a; Kurcz, 2005, 2011; Fromkin i Ratner, 2005).

### **2.2.1 Model Fromkin-Garreta**

Pierwsze współczesne (psycholingwistyczne) modele produkcji mowy powstawały głównie w oparciu o dane płynące z gromadzonych przykładów błędów mowy (Harley, 2001; Levelt, 1999a). Wynikało to z założenia, że przejęzyczenia, poprzez swoją regularność mogą odsłonić mechanizmy odpowiedzialne za kodowanie przekazu werbalnego (Fry, 1969).

Analizując spontanicznie popełniane pomyłki w mówieniu, dostrzec można różne ograniczenia jakim podlegają wadliwe realizacje zaplanowanej wypowiedzi. Na przykład błędy zamiany całych słów (substytucje), w których zamiast docelowego słowa wypowiedzane jest inne (np. „noc” zamiast „dzień”) podlegają ograniczeniu do jednej kategorii syntaktycznej. Z kolei błędy polegające na zamianie sylab między wyrazami, zachodzą między blisko sąsiadującymi ze sobą słowami, niezależnie od kategorii syntaktycznej tych słów, przy jednoczesnym ograniczeniu potencjalnych błędów do wymiany sylab o podobnej strukturze. Takie zjawiska pozwoliły wysnuć hipotezę, że błędy te musiały powstać na różnych etapach procesu mówienia, a tym samym, że ma on złożoną, kilkupoziomową organizację.

Na przykład Fay i Cutler (1977) analizując błędy zamiany całych słów (substytucje) wysunęli postulat złożonej struktury słownika umysłowego. Ich zdaniem poszczególne jednostki leksykonu mentalnego łączą nie tylko relacje semantyczne, ale również podobieństwo formy. Poparciem tej hipotezy są liczne przykłady przejęzyczeń, w których między słowem docelowym a pomyłkowym istnieje wyraźne podobieństwo fonologiczne, przy braku jakichkolwiek związków semantycznych. Według Faya i Cutler (1977) gdyby

słownik umysłowy zorganizowany był jedynie według relacji semantycznych, tego typu błędy nie powinny się zdarzać.

Jednym z pierwszych modeli procesu mówienia opartym na kompleksowej analizie różnorodnych błędów mowy był schemat zaproponowany przez Fromkin (1973). Badaczka zaobserwowała, że stosowany w językoznawstwie podział na różne jednostki strukturalne języka (np. fonemy, morfemy itd.) ma swoje odzwierciedlenie w błędach mowy. Na przykład błędy występujące w przejęzyczeniach dotyczą jednostek z każdego poziomu organizacji języka, od cech dystynktywnych, po całe frazy. Jednocześnie wadliwa realizacja jakiegoś elementu języka, może wcale nie zaburzyć prawidłowej budowy pozostałej części wypowiedzi. Oznaczałoby to, że kodowanie poszczególnych aspektów komunikatu językowego odbywa się w pewnej niezależności od siebie. Doprowadziło to Fromkin (1973) do wniosku, że elementy wypowiedzi ulegające błędnym transformacjom nie tylko dokładnie odpowiadają jednostkom języka, ale prawdopodobnie odzwierciedlają jednostki i etapy programowania wypowiedzi. Jest to założenie o wielopoziomowej i wieloetapowej złożoności procesu mówienia (podobnie jak koncepcji płaszczyzn mowy Wygotskiego), które legło u podłoża późniejszych modeli. Generowanie wypowiedzi jest tu ujęte jako ciąg procesów odpowiedzialnych za kodowanie odrębnych jednostek językowych. Do tych procesów należą kodowanie semantyczne, leksykalne, morfologiczne, fonologiczne oraz artykulacja.

Propozycję Fromkin (1973) rozwinął w swoich pracach Garrett, który również opracował model bazujący na szczegółowej analizie przejęzyczeń i ujmujący proces mówienia, jako serię następujących po sobie niezależnych poziomów kodowania wypowiedzi językowej (Garrett, 1976; Harley, 2001). Twierdził on, że dane z analizy błędów mowy odzwierciedlają zmiany jakie przechodzi zdanie w procesie mówienia począwszy od pojęciowej (niejęzykowej) reprezentacji przekazu (*message level*) po jego reprezentację fonologiczną, dostarczającą bezpośrednich instrukcji wykonawczych dla aparatu artykulacyjnego. Według niego planowanie wypowiedzi odbywa się na przynajmniej dwóch odrębnych poziomach przetwarzania (modułach), z których pierwszy (*functional level*) odpowiada za wybór słownictwa i wygenerowanie ramy syntaktycznej dla całej wypowiedzi, drugi natomiast (*positional level*) – za sekwencyjną organizację morfemów i fonemów w ciągu wypowiedzi (Garrett, 1976).

Poziomem wyjściowym w modelu i zarazem pierwszym etapem procesu mówienia jest poziom przekazu. Powstaje na nim pojęciowa reprezentacja treści zdania, wynikająca z intencji komunikacyjnej nadawcy komunikatu. Na tym etapie nadawca decyduje o tym, co – i w jaki sposób – zamierza powiedzieć (Harley, 2001). Garrett nie opisywał szczegółowo

charakterystyki tego poziomu, zakładając jedynie że generuje on informację wejściową dla kolejnych poziomów odpowiedzialnych za językowe programowanie wypowiedzi. Zakładał jednak, że informacje obecne na tym poziomie są oparte na obecnym stanie percepcyjnym i uczuciowym mówcy oraz jego ogólnej znajomości świata (Garrett, 1993).

Właściwe kodowanie rozpoczyna się na poziomie funkcjonalnym. Na tym etapie ze słownika umysłowego wybierane są jednostki leksykalne, odpowiadające znaczeniowo treści wygenerowanej na wcześniejszym poziomie. Proces selekcji odbywa się zgodnie z regułami semantycznymi, aby użyte wyrazy treściowe (*content words*), jak najlepiej oddawały powzięty zamiar komunikacyjny. Jednocześnie z wybranych jednostek leksykalnych tworzona jest rama syntaktyczna zdania, w oparciu o ich kategorie gramatyczne (rzeczownik, czasownik, przymiotnik itd.) i funkcje składniowe (podmiot, orzeczenie, dopełnienie itd.). Procesy operujące na poziomie funkcjonalnym działają więc zgodnie z regułami semantycznymi i syntaktycznymi, prowadząc do wyboru jednostek leksykalnych i przypisania im odpowiedniej roli w zdaniu, bez ostatecznego określenia ich szyku (Garrett 1976). Na tym etapie generowane zdanie ma jeszcze postać kanoniczną i jest reprezentowane przez abstrakcyjną strukturę predykatowo-argumentową (Kurcz, 2005).

Według Garretta (1976) taka organizacja poziomu funkcjonalnego wyjaśnia powstawanie błędów zamiany miejscami całych wyrazów w zdaniu (*word exchanges*). Są one najczęściej ograniczone do wspólnej kategorii gramatycznej, a ich przestawienie w planie syntagmatycznym pociąga za sobą większe zmiany w strukturze powierzchniowej zdania, niż tylko samo ich przesunięcie i zamiana pozycji. Na przykład w błędnym zdaniu (Bręński, 2013, s. 92) „brat może przejąć jej funkcję” (docelowo zamierzone zdanie: ona może przejąć funkcję brata) poza samą zamianą miejsca, transformacji uległa również:

1. forma wyrazu „brat” (wymuszona zmianą przypadku z dopełniacza na mianownik),
2. zmieniony został zaimek osobowy „ona” na zaimek dzierżawczy „jej”
3. dopasowany został szyk wyrazów tak, aby wyraz „funkcja” w sposób właściwy składniowo i stylistycznie połączył się z zaimkiem „jej”.

Według Garretta tego typu przykłady dowodzą, że zamiana całych wyrazów może zajść na głębszym, bardziej abstrakcyjnym poziomie, gdzie nie został jeszcze doprecyzowany ich szyk ani forma fonologiczna. W ten sposób błąd powstały na poziomie funkcjonalnym, o ile nie prowadzi do konstrukcji łamiącej reguły języka, może zostać poddany dalszemu kodowaniu, prowadząc do językowo poprawnej, choć niezgodnej z intencją nadawcy, wypowiedzi.

W związku z tym, że na poziomie funkcjonalnym podstawowymi elementami kodowania są jednostki leksykalne, Garrett lokuje na nim również źródło powstawania

przejęzyczeń o charakterze substytucji całych słów oraz tak zwane zmieszania (*blends*). Te pierwsze cechują się tym, że słowo docelowe i pomyłka pochodzą z tej samej kategorii gramatycznej (rzeczownik wymienia się na rzeczownik, czasownik – na czasownik itd.), a także łączą je zazwyczaj jakieś relacje semantyczne lub skojarzeniowe. Błędy zmieszania z kolei polegają na wypowiedzeniu jednocześnie dwóch wyrazów (np. „marsonka” zamiast „marynarka” lub „garsonka”). W tego typu błędach relacje semantyczne są jeszcze silniejsze i najczęściej mają charakter synonimii (Bręński, 2013; MacKay, 1980).

Odmienne właściwości wykazują natomiast pomyłki, które Garrett (1976) lokuje na kolejnym poziomie procesu mówienia. Nie zachowują one zazwyczaj zgodności kategorii syntaktycznej i zachodzą w obrębie tej samej frazy, między bezpośrednio sąsiadującymi ze sobą wyrazami. Ich powstawanie ograniczone jest jednak do zamiany lub przesunięcia jednostek na tym samym poziomie strukturalnym, niższym niż poziom całego wyrazu (np. morfem, sylaba itp.). Zdaniem Garretta, powstawanie tego rodzaju błędów można przypisać zakłóceniom procesów kodowania fonologicznego działającym na poziomie pozycyjnym. Jest to etap generowania wypowiedzi, na którym fonologiczne reprezentacje słów są układane w ściśle określonym szyku wyznaczanym przez wyrazy funkcyjne (Garrett, 1976). Na poziomie pozycyjnym zdanie jest doprecyzowane pod względem szyku generowanej wypowiedzi oraz swojej abstrakcyjnej formy dźwiękowej (fonologicznej). O tym, że ostateczny kształt fonetyczny generowana wypowiedź otrzymuje dopiero w kolejnym etapie procesu mówienia, świadczą akomodacje (*accommodations*) (Garrett, 1980), polegające na tym, że błędnie przemieszczone morfemy lub sylaby dopasowują się pod względem fonetycznej formy do nowego otoczenia (nowego miejsca w szyku) zgodnie z regułami fonotaktyki i podczas artykulacji podlegają zwyczajnym przekształceniom fonetycznym, jak upodobnienia czy zmiękczenia (Bręński, 2015; Garrett, 1980, Harley, 2001).

Według Harleya (2001) zaobserwowane przez Garretta różnice w przejęzyczeniach między błędami popełnionymi na (całych) wyrazach treściowych, a gramatycznych strukturach funkcyjnych (morfemy fleksyjne i wyrazy funkcyjne) świadczą o tym, że są one przetwarzane na różnych poziomach procesu mówienia, co uzasadnia rozdzielenie kodowania językowego na dwa osobne poziomy.

W modelu Garretta, po poziomie funkcjonalnym i pozycyjnym następują jeszcze dwa osobne etapy programowania wypowiedzi. Pierwszy z nich, który na wejściu przetwarza informację z poziomu pozycyjnego, tworzy fonetyczną reprezentację tworzonego zdania. Na tym etapie zachodzi zjawisko akomodacji pozwalające na realizację przejęzyczeń, które otrzymują dopuszczalną przez system fonotaktyczny formę dźwiękową. Poziom ten jest więc

odpowiedzialny za ostateczne przygotowanie zdania do artykulacji. Jest to również ostatni poziom wewnętrznego programowania wypowiedzi, na którym mogą według Garretta (1976) wystąpić jeszcze błędy polegające na pominięciu jakiejś głoski lub grupy głosek (*simple and complex sound deletions*). Poziom reprezentacji fonetycznej przygotowuje bezpośrednie instrukcje dla artykulatorów. W procesie kodowania motorycznego ruchu aparatu mowy tworzą ostatni poziom reprezentacji, nazywany reprezentacją artykulacyjną (Garrett, 1993). Jest to zewnętrzny etap procesu mówienia. Powstałe w wyniku ruchów artykulacyjnych zmiany ciśnienia powietrza tworzą falę akustyczną, rozpoznawaną przez odbiorcę jako dźwięki mowy. Zdaniem Garretta, podczas kodowania reprezentacji artykulacyjnej zachodzą zjawiska obserwowane w tzw. „łamaczach języka” (*tongue twisters*). Schematyczne ujęcie całego modelu prezentuje rysunek 2.

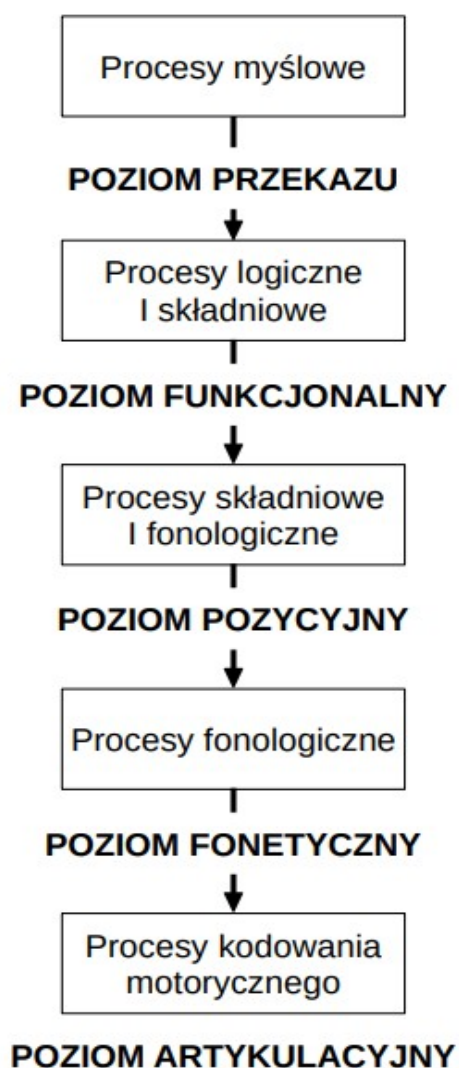
Zaproponowany przez Fromkin (1973) i rozwinięty przez Garretta (1976, 1980) model w całości oparty został na analizie spontanicznych przejęzyczeń. W założeniu miał jednak opisywać normalny proces mówienia. Za jego trafnością przemawiać może materiał empiryczny – przejęzyczenia zbierane w naturalnych, codziennych sytuacjach lepiej odzwierciedlają procesy zachodzące w mowie zwykłych użytkowników języka, niż sztuczne zadania eksperymentalne ograniczające spontaniczność procesu mówienia. Jest to zarazem słaba strona tej metody pozyskiwania danych, gdyż nie pozwala na kontrolowanie warunków w których przebiega obserwacja. Choć badacz może wykluczyć metodę anegdoty i do pewnego stopnia wytrenować umiejętność „wychwytywania” przejęzyczeń z mowy innych osób, w sytuacji zbierania materiału sam jest zazwyczaj zaangażowany w komunikację (jest uczestnikiem danej sytuacji komunikacyjnej), przez co uzyskane dane mogą nie być w pełni reprezentatywne np. pod względem dystrybucji różnych typów przejęzyczeń (por. Cutler i Fay, 1978; Harley, 2001; Harley i MacAndrew, 2001). Między innymi z tego względu konieczne stało się opracowanie innych modeli, które uwzględniałyby różne źródła danych na temat procesu mówienia.

Na bazie modelu Fromkin i Garretta oraz jemu podobnych lub pokrewnych wyrosły dwie najbardziej wpływowe współcześnie propozycje teoretycznego ujęcia produkcji mowy. Levelt (1989) opisał przebieg procesu mówienia jako szereg transformacji myśli na kod językowy, dokonujących się w wyspecjalizowanych modułach odpowiadających za kolejne przekształcenia komunikatu. Była to propozycja najbardziej kompletna spośród ówczesnych modeli, obejmująca proces mówienia od sformułowania intencji do jej wyartykułowania oraz uwzględniająca szerokie spektrum danych (np. eksperymentalnych), wychodząc tym samym poza analizę błędów mowy. Dell (1986) natomiast pozostając w tradycji analizy danych z

korpusów przejęzyczeń opracował pierwszy model obliczeniowy produkcji mowy, którego częściowa formalizacja pozwalała na predykcję częstości pewnych typów błędów. Modele te stały się konkurencyjnymi wyjaśnieniami przebiegu procesu mówienia i stymulowały liczne badania mające na celu dostarczenie danych rozstrzygających między tymi modelami (Levelt, 1993; 1999a; por. Roelofs, 2020a; Nozari, 2020).

## Rysunek 2

*Poziomy produkcji mowy według Garretta*



Źródło: Garrett (1976; 1980)

## 2.2.2 Model produkcji mowy według Levelta (1989)

Levelt (1989) zauważył, że przy tworzeniu wcześniejszych modeli procesu mówienia zbyt dużo uwagi przykładano do przejęzyczeń, nie biorąc pod uwagę innych źródeł danych na temat procesu mówienia. Chcąc wypełnić tę lukę zaproponował własną oryginalną koncepcję procesu mówienia uwzględniającą bogaty zestaw danych eksperymentalnych.

Model opracowany przez Levelta w publikacji *Speaking: From Intention to Articulation* z 1989 roku wpisuje się w popularną wówczas komputerową metaforę umysłu. Po pierwsze zakłada on, że operacje umysłowe, które doprowadzają do wypowiedzenia określonego słowa lub zdania przebiegają na odrębnych poziomach, funkcjonujących jak moduły<sup>36</sup> złożonego układu. Struktura tego procesu ma charakter hierarchiczny, to znaczy że informacja przetwarzana jest etapowo, na odrębnych poziomach, w których przepływ informacji jest jednokierunkowy. Inaczej mówiąc przetwarzanie informacji, która ma przyjąć formę wypowiedzi głosowej, przebiega sekwencyjnie i w sposób izolowany, co oznacza, że na danym poziomie właściwe dla niego operacje rozpoczynają się dopiero po zakończeniu pracy na poziomie wcześniejszym i hierarchicznie wyższym. Poszczególne komponenty kodowania informacji zachowują swoją hermetyczność (*encapsulation*), co oznacza że przepływ informacji między nimi nie jest swobodny, lecz ograniczony do połączeń między „wyjściem” z jednego modułu i „wejściem” do kolejnego. Procesy zachodzące w poszczególnych komponentach są także wysoce zautomatyzowane<sup>37</sup> (*automaticity*) oraz niedostępne innym świadomym procesom poznawczym w tym sensie, że przekonania, pragnienia, czy cele podmiotu (nadawcy) nie są w stanie zmienić przepływu danych w modułach (Wróbel, 2007). W późniejszym czasie model uległ pewnym modyfikacjom (np. dodane zostało niewystępujące wcześniej pojęcie sylabariusza<sup>38</sup>, czyli umysłowego repozytorium wzorców motoryczno-akustycznych) związanym m. in. z jego formalizacją i dopasowaniem do danych eksperymentalnych (Levelt i in., 1999; Levelt, 2001b). W tym

---

36 Levelt (1989) zdecydował się nie nazywać wyróżnionych przez siebie komponentów przetwarzania informacji modułami, gdyż przypisuje im cechy, które tylko częściowo pokrywają się z właściwościami modułów postulowanymi przez kognitywistów. Dyskusja na temat modularności umysłu ma swoją już kilkudziesięcioletnią tradycję w naukach o poznaniu i przez ten czas nie udało się wypracować ani uwspólnionej definicji modułu ani też określić jakie ogólne właściwości można przypisać różnym modułom (por. Miłkowski, 2007; Wróbel 2007). Z tego względu cechy, jakie Levelt (1989) przypisuje wyróżnionym przez siebie komponentom produkcji mowy, tj. automatyzm (*automaticity*), hermetyczność informacyjna (*informational encapsulation*), poznawcza nieprzenikliwość (*cognitive impenetrability*), pozwalają ulokować jego model bliżej teorii modularnych niż koneksjonistycznych. Poza tym, niezależnie od unikania używania terminu „moduł”, inni badacze zaliczają model Levelta do grupy teorii modularnych, nawet w jego późniejszej, sieciowej wersji (por. sekcja komentarzy do artykułu Levelta i in. 1999).

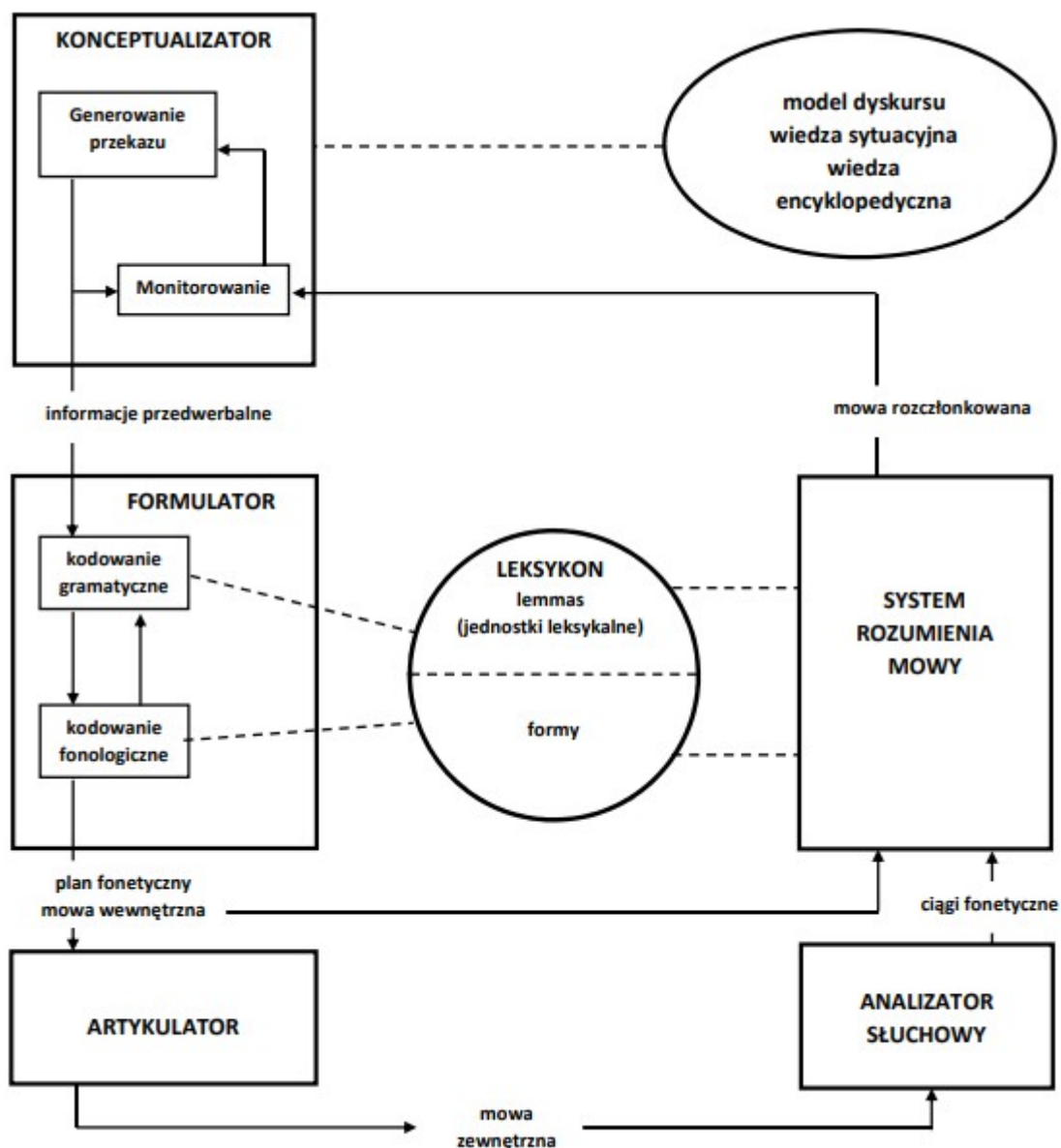
37 Problem kontroli poznawczej nad procesem mówienia zostanie poruszony w dalszej części niniejszej pracy.

38 Levelt (1989) używał pierwotnie określenia „inwentarz sylab” (*inventory of syllables*), który miał odpowiadać hipotetycznemu zbiorowi gotowych programów ruchowych odpowiadających często realizowanym w mówieniu planom fonetycznym.

miejszu opisany zostanie pierwotny model zaproponowany przez Levelta w 1989 (graficzną reprezentację modelu przedstawia rysunek 3), natomiast wprowadzone do niego zmiany zostaną opisane w dalszej części rozdziału i potraktowane jako osobny model.

### Rysunek 3

Model produkcji mowy według Levelta



Źródło: Levelt (1989), tłumaczenie za: Frydrychowicz (1999)

W omawianym modelu tworzenie wypowiedzi językowej rozpoczyna się w komponencie nazywanym przez Levelta (1989) konceptualizatorem<sup>39</sup>. Jest on odpowiedzialny za wygenerowanie przedwerbalnej (konceptualnej) reprezentacji treści wypowiedzi oraz za monitorowanie i ewaluację jej językowej realizacji na dalszych etapach produkcji mowy. Podczas procesów konceptualizacji ustalana jest intencja komunikacyjna nadawcy oraz informacje, które będą w stanie ją wyrazić. Na tym etapie informacje te nie mają jeszcze postaci językowej, lecz występują w postaci pewnych klas pojęciowych, a konceptualizator czerpie je z posiadanej przez osobę mówiącą wiedzy encyklopedycznej, zapisanej w pamięci długotrwałej, a także z dostępnych danych kontekstowych na temat sytuacji komunikacyjnej, w jakiej się znajduje. Efektem końcowym procesu konceptualizacji, informacją wejściową dla kolejnego modułu (formulatora) jest przekaz przedwerbalny (*preverbal message*), będący pojęciową (semantyczną) reprezentacją powstającej wypowiedzi, która odnosi się do jakiegoś stanu rzeczy. Choć taki stan rzeczy może być reprezentowany w umyśle w dowolny sposób (w kodzie wizualnym, werbalnym, sensomotorycznym), to reprezentujący je w procesie mówienia przekaz przedwerbalny ma postać abstrakcyjnego sądu o strukturze predykatowo-argumentowej (*proposition*). Ponadto generowany przekaz zawiera pozostałe informacje o strukturze wypowiedzi dotyczące np. jej modalności, tematu-rematu, a także tych charakterystyk, które w danym języku mogą mieć swoją specyficzną realizację powierzchniową, np. charakterystyka temporalna (por. Grzegorzyczkowa, 2001).

Według Levelta (1989) kodowanie tych składników odbywa się dwuetapowo w procesach makro- i mikroplanowania. Makroplanowanie (*macroplanning*) odpowiada za opracowanie intencji komunikacyjnej (modalność wypowiedzi) w postaci celów i podcelów, którym przypisane zostają odpowiednie informacje. Wyrażenie tych informacji w formie językowej wypowiedzi, dzięki językowemu kodowaniu na kolejnych etapach zaspokaja cele opracowane w czasie makroplanowania i umożliwia realizację intencji komunikacyjnej. W drugim kroku, tj. podczas mikroplanowania (*microplanning*) ustalany jest temat (i remat) wypowiedzi, a wybrane podczas makroplanowania informacje zostają wpisane w odpowiednie miejsca strukturalnej ramy propozycjonalnej (predykatowo-argumentowej). Levelt dopuszcza, że ta wysoce abstrakcyjna i uniwersalna forma reprezentacji może częściowo różnić się u osób mówiących różnymi językami ze względu na różnice występujące między językami w tym, jakie kategorie gramatyczne są wyrażane w strukturze powierzchniowej wypowiedzi (dotyczy to np. deklinacji i koniugacji). Przypomnijmy, że według Hjelmsleva (1979) istniejąca pomiędzy formą wyrażania i formą treści

---

39 Model opisywany jest - podobnie jak w pracy Kurcz (2005) - zgodnie z tłumaczeniem na język polski wykonanym przez Frydrychowicza (1999).

współzależność, nie ma charakteru bezwzględnie symetrycznego. Trudno więc rozstrzygnąć, jaką ostatecznie postać przyjmuje zestaw informacji opuszczający moduł konceptualizacji, tym bardziej, że – jak zauważa Harley (2001) – konceptualizacja i format w jakim zapisany jest przekaz przedwerbalny są najsłabiej poznanymi aspektami procesu mówienia. Ponadto dobór informacji do bezpośredniego wyrażenia w wypowiedzi jest wysoce uzależniony od kontekstu rozmowy i niektóre informacje są traktowane jako domyślne, to znaczy że możliwe jest ich wywnioskowanie z danej wypowiedzi nie wprost. Informacje takie nie mają charakteru semantycznego (nie są przekazywane w kodzie językowym), lecz pragmatyczny (Grzegorzczkova, 2001). Oznacza to, że nie stanowią one składników przekazu, ale muszą być brane pod uwagę przez nadawcę podczas konceptualizacji.

W momencie, kiedy wygenerowany zostaje przekaz przedwerbalny, konceptualizator kończy przetwarzanie danej wypowiedzi i zostaje ona poddana operacjom formulatora. Zasadniczym sposobem działania tego komponentu jest przetłumaczenie przekazu o strukturze pojęciowej na kod językowy (werbalna postać wypowiedzi). Proces ten, podobnie jak w poprzednim module, odbywa się w dwuetapowo w postaci kodowania gramatycznego (*grammatical encoding*) oraz kodowania fonologicznego (*phonological encoding*). W pierwszym kroku koder gramatyczny uzyskuje dostęp do jednostek leksykonu umysłowego (tzw. lemmas), które zostają umieszczone w ramie syntaktycznej, tworząc w ten sposób strukturę powierzchniową zdania. Rama syntaktyczna generowana jest na podstawie wskazówek dostarczonych do formulatora przez strukturę przekazu przedwerbalnego. Umieszczane w niej lemmy zawierają informacje o charakterze:

- a) semantycznym – każde przywołane z pamięci (podczas konceptualizacji) pojęcie otrzymuje przypisaną mu odpowiednią lemmę (znaczeniowo zgodną z danym pojęciem);
- b) syntaktycznym – każda lemma otrzymuje właściwe miejsce w sekwencyjnej strukturze zdania zgodnie z porządkiem wyznaczanym przez jej funkcję i kategorię gramatyczną.

Wydobycie właściwej lemmy kończy selekcję leksykalną (*lexical selection*, Levelt, 2001b), polegającą na zawężeniu wzbudzonej treści pojęciowej do znaczenia jednego konkretnego słowa. Dzięki posiadanym przez lemmy wskazówkom haseł leksykalnych (leksemów) możliwe jest rozpoczęcie kodowania fonologicznego (drugi etap pracy formulatora). Proces ten polega na wydobyciu z leksykonu umysłowego odpowiednich form leksykalnych, które charakteryzują poszczególne lemmy pod względem morfologicznym i fonologicznym. Innymi słowy każdej lemmy przypisane zostają morfemy leksykalne i

fleksyjne, które wyznaczają porządek ciągu fonemów tworzących daną wypowiedź (wyraz, zdanie itd.).

Metaforycznie można powiedzieć, że jednostki leksykalne wyrażające treść i kategorię gramatyczną słowa mają swoje miejsce w „leksykonie lemm”, a leksemy, czyli morfofonologiczne postaci słów, znajdują się w „leksykonie form” (Levelt, 1989). Lemmy zorganizowane są pod względem wzajemnych relacji semantycznych, natomiast leksemy „sąsiadują” ze sobą ze względu na podobieństwo formalne (podobna struktura fonologiczna) oraz częstość używania w mowie (Jescheniak i Levelt, 1994; Levelt, 2001b).

Produktem końcowym kodowania fonologicznego i formą informacji na wyjściu z formulatora jest fonetyczny lub artykulacyjny plan. Levelt utożsamia go z pojęciem mowy wewnętrznej<sup>40</sup> (*internal speech*), gdyż nadal jest to nieuzewnętrzniony poziom reprezentacji, który zawiera informację o tym, w jaki sposób dane słowa lub zdania należy wypowiedzieć. Są to bezpośrednie wskazówki dla działania modułu artykulatora.

Artykulator, czyli ostatni moduł procesu mówienia odpowiedzialny jest za motoryczną realizację fonetycznego planu, jaki powstał w wyniku operacji wykonanych przez formulator. Można metaforycznie powiedzieć, że pełni on rolę interfejsu tłumaczącego kod językowy (fonetyczny) na impulsy nerwowe kierowane do efektorów narządów mowy. Przed motorycznym wykonaniem planu fonetycznego może on być tymczasowo buforowany. Według Levelta (1989) działanie bufora artykulacyjnego niweluje różnice między szybkością formułowania i szybkością artykulacji, pozwalając tym samym zachować płynność ruchów artykulacyjnych dla całej wypowiedzi (lub jej części) pomimo zmiennego tempa opracowywania informacji językowej na poziomie formulatora. Inaczej mówiąc artykulator otrzymując niepełny plan fonetyczny wypowiedzi z formulatora wstrzymuje jego realizację do czasu uzyskania pełniejszej porcji informacji. Zdaniem Levelta (1989), taką minimalną jednostką gotową do wypowiedzenia jest wyraz fonologiczny (*phonological word*)<sup>41</sup>. Bufor artykulacyjny służy także do chwilowego przechowywania gotowego już planu fonetycznego i przygotowania narządów mowy do artykulacji nowej porcji informacji, jeszcze podczas realizacji wcześniejszych instrukcji artykulacyjnych. Skurcze mięśni powodujące ruch tzw. aparatu mowy modyfikują strumień powietrza przepływający przez układ oddechowy, wprowadzając go w drgania. Powstająca w ten sposób fala akustyczna odbierana jest drogą słuchową zarówno przez nadawcę, jak i odbiorcę jako dźwięki mowy. Tym samym droga

---

40 Levelt używa terminu „mowa wewnętrzna” (*inner speech*) w znaczeniu innym niż Wygotski (1934/1989), gdyż utożsamia ją z bezgłośnym przetwarzaniem konkretnych słów i fraz.

41 Wyraz fonologiczny – to „ciąg fonemów zawartych między dwiema potencjalnymi pauzami z jedną (i tylko jedną) sylabą – nosicielem akcentu głównego” (Polański, 2003, s. 646).

wewnętrznych operacji prowadzących do wygenerowania wypowiedzi zostaje zakończona, a wypowiedź staje się elementem mowy zewnętrznej (*overt speech*).

Levelt (1989) uwzględnia w swoim modelu także procesy związane z odbiorem mowy, jednak nie stanowią one istotnej części jego teorii. Włączenie ich do modelu produkcji mowy jest o tyle zasadne, że opisany model czyni również pewne założenia dotyczące monitorowania powstającej wypowiedzi, dzięki któremu możliwa jest jej ewaluacja pod względem językowej poprawności oraz zgodności z założonym planem i wzbudzoną intencją. Modułem odpowiedzialnym za monitorowanie procesu mówienia jest konceptualizator, który ma bezpośredni dostęp do intencji stojącej za wygenerowaną wypowiedzią, dzięki czemu może dokonać porównania wygenerowanych podczas makroplanowania celów z ich językową realizacją. Według Levelta, konceptualizator korzysta w tym celu z informacji dostarczonych przez system odbioru mowy (*speech comprehension system*), który ma dostęp zarówno do wypowiedzi już wyartykułowanej (odebranej drogą słuchową), ale także do mowy wewnętrznej występującej w postaci planu fonetycznego. Dzięki temu możliwe jest wykrycie nieścisłości i pomyłek nie tylko po ich wypowiedzeniu, ale zanim jeszcze trafią do artykulatora. To z kolei umożliwia poprawienie tych błędów jeszcze na poziomie bufora artykulacyjnego, zanim zdążą się uzewnętrznić. Bardziej szczegółowo zagadnienie monitorowania i kontroli procesu mówienia zostanie omówione w rozdziale 3.

### **2.2.3 Koneksjonistyczny model Della (1986)**

Jeden z pierwszych koneksjonistycznych modeli procesu mówienia został zaproponowany przez Della. Badacz ten pracował nad swoją teorią od wczesnych lat 80-tych XX w. i rozwijał ją przez kolejne trzy dekady. W zasadzie być może bardziej adekwatne byłoby mówienie o modelach produkcji mowy, gdyż szczegółowe propozycje tego w jaki sposób można modelować proces mówienia lub jego poszczególne etapy, zmieniały się w kolejnych publikacjach. Jednocześnie niezmienny pozostał trzon teorii Della, która zakłada sieciowy charakter powiązań między poszczególnymi elementami i regułami umysłowej reprezentacji języka, biorącymi udział w generowaniu wypowiedzi. Poniżej zaprezentowane zostaną właśnie te podstawowe i trwałe założenia teorii Della (1986), opracowane w połowie lat 80-tych XX w. i opublikowane w czasopiśmie *Psychological Review*. Z tych założeń wyprowadzane były w kolejnych latach szczegółowe twierdzenia dotyczące przebiegu poszczególnych etapów generowania wypowiedzi (kodowanie fonologiczne, morfologiczne, gramatyczne) sprawdzane następnie w różnych symulacjach opartych o sieci neuronowe.

Punktem wyjścia dla sformułowania teorii było dla Della (1986) założenie, że model wytwarzania mowy powinien wyjaśniać zarówno proces tworzenia poprawnej (zgodnej z regułami języka i intencją komunikacyjną nadawcy) wypowiedzi, jak i sformułowań klasyfikowanych jako pomyłki, czy też przejęzyczenia. Podobnie jak Fry (1969), Fromkin (1973), czy Garrett (1976, 1980), Dell uważał, że analiza spontanicznych przejęzyczeń dostarcza danych na temat jednostek, jakie biorą udział w generowaniu wypowiedzi. Dlatego zaproponowany przez niego model nie tylko ma wyjaśniać dlaczego powstają określone przejęzyczenia, ale także symulować proces mówienia, przewidując szansę wygenerowania prawidłowej wypowiedzi lub popełnienia przez osobę mówiącą – czy raczej wygenerowania przez system – błędu.

Według Della (1986) proces generowania wypowiedzi<sup>42</sup> polega na tłumaczeniu semantycznej reprezentacji tego, co ma zostać powiedziane na jej fonetyczną formę, która wyznacza ruchy aparatu artykulacyjnego. Proces ten przebiega na kilku poziomach przetwarzania informacji językowej jednocześnie (zgodnie z klasyfikacją Fry'a, 1969), a zdanie, które zamierza wypowiedzieć osoba mówiąca jest na każdym z tych poziomów reprezentowane poprzez inne jednostki i reguły języka (pojęcia, wyrazy<sup>43</sup>, morfemy, fonemy). Wszystkie poziomy generowania wypowiedzi tworzą hierarchiczną sieć jednostek i reguł językowych (*network of linguistic rules and units*). Podczas mówienia, dzięki mechanizmowi rozprzestrzeniającej się po sieci aktywacji, wzbudzone (aktywowane) są poszczególne węzły tej sieci, reprezentujące różne elementy systemu językowego (jednostki i reguły). Aktywacja między węzłami rozprzestrzenia się w sposób kaskadowy oraz dwukierunkowy, a zatem między poziomami przetwarzania – odmiennie niż w modelu Levelta (1989) – zachodzi interakcja, a przetwarzanie generowanej wypowiedzi odbywa się na kilku poziomach jednocześnie. Szlaki po jakich rozprzestrzenia się aktywacja łączą węzły w relacjach pionowych (między hierarchicznie zorganizowanymi poziomami) oraz w relacjach poziomych, łącząc jednostki kodowania (pojęcia, lemmy, morfemy, fonemy) ze względu na jakiś rodzaj podobieństwa, czy współwystępowania w użyciu (relacje semantyczne, skojarzenia, częstość użycia, podobna struktura itp.). W związku z tym, że aktywacja rozprzestrzenia się wielokierunkowo, poziom aktywacji danego węzła jest efektem sumowania się aktywacji pochodzącej z różnych źródeł (od innych węzłów). Węzły o najwyższym poziomie aktywacji determinują to, w jaki sposób reprezentowana jest wypowiedź na danym poziomie kodowania (np. jaki wyraz zostanie wybrany na określenie

42 Prototypową wypowiedzią jest dla Della (1986) zdanie, dlatego podobnie jak w modelu Garretta (1975) opisuje on w jaki sposób przebiega wytwarzanie zdania (*sentence production*).

43 Podobnie jak Levelt (1989), Dell (1986) za podstawową jednostkę leksykalną uznaje lemmę, czyli hasłową (słownikową) formę wyrazu.

danego pojęcia: mebel do siedzenia bez oparcia może być określony jako „stołek” lub „taboret”), co – w przypadku zwiększonej aktywacji niewłaściwego węzła – może skutkować wystąpieniem błędu, który da się słyszeć jako przejęzyczenie (np. wybranie wyrazu „fotel” zamiast „stołek” albo fonemu /r/ zamiast /t/ w nagłosie wyrazu „taboret”, co utworzy przejęzyczenie „raboret”). Według Della (1986) popełnianie pomyłek podczas mówienia jest nieuniknione gdyż język jest systemem otwartym i produktywnym, a zatem obsługująca go sieć leksykalna (*lexical network*) musi do pewnego stopnia pozostać elastyczna i otwarta na nowe konstrukcje, a tym samym pozwolić na pewien tolerowalny odsetek błędów. Aby minimalizować ryzyko popełnienia błędu model Della oprócz rozprzestrzeniania się aktywacji po sieci i jej sumowania w węzłach, zakłada również procesy zanikania (*decay*) aktywacji w węzłach<sup>44</sup>. Jak wspomniano, na każdym z poziomów przetwarzania generowana wypowiedź otrzymuje swoją specyficzną reprezentację, zgodnie z charakterystyką jednostek i reguł przynależnych do danego poziomu. Utworzenie reprezentacji wypowiedzi na danym poziomie odbywa się poprzez wygenerowanie abstrakcyjnej ramy taktycznej (*tactical frame*), której „okienka” (*slots*) wypełniane są jednostkami strukturalnymi języka, charakterystycznymi dla danego poziomu. Jednostki te przechowywane są w pamięci długotrwałej, w tzw. leksykonie<sup>45</sup>, traktowanym przez Della (1986) nie jako osobny moduł, z którego można wydobyć dany element (np. lemmę, morfem, czy fonem), ale jako część sieci (poszczególne węzły sieci), której aktywacja przebiega od jednostek strukturalnie „większych” (pojęcia, lemmy) do bardziej podstawowych (morfemy, fonemy). W związku z tym, że chodzi tu o relacje sieciowe, charakter powiązań między węzłami jest wzajemny. Oznacza to, że zarówno wzbudzenie elementu na wyższym poziomie aktywuje powiązane z nim elementy na niższym poziomie, jak i na odwrót: jednostka niższego rzędu przesyła część aktywacji do skojarzonych z nią jednostek na wyższym poziomie. Oprócz jednostek budujących ramę taktyczną, w sieciowych relacjach na każdym z poziomów kodowania pozostają także reguły tworzące tę ramę (*generative rules*), które Dell (1986) odróżnia wyraźnie od leksykonu. Ogólny kształt ramy taktycznej wyznacza porządek wypełniania „okienek” jednostkami kodowania, co jest zgodne z systemowym ujęciem języka, w którym jednostki wyższego rzędu są strukturujące dla jednostek niższego rzędu (por. rozdział 1).

44 W pierwotnym modelu Dell (1986) zakładał, że połączenia między węzłami mają raczej charakter wzmacniający niż blokujący, w związku z czym przewidywana była przewaga procesów pobudzenia nad hamowaniem, a dezaktywacja węzłów następować miała przez pasywne zanikanie aktywacji. W późniejszym modelu DoL (*division of labor*) symulującym przejście z poziomu semantycznego do kodowania gramatycznego w generowaniu prostych, dwuwyrzowych zdań Dell i in., (2008) przypisali połączeniom między węzłami wagi, które mogły być dodatnie lub ujemne.

45 W przeciwieństwie do Levelta (1989), który leksykonem określał zbiór lemm i leksemów (słowoform) – Dell (1986) używa terminu „leksykon” w odniesieniu do każdego umysłowego repozytorium jednostek budujących wypowiedź na wszystkich poziomach organizacji procesu mówienia.

Pierwszy poziom wytwarzania wypowiedzi związany jest z skonstruowaniem jej semantycznej reprezentacji. Składa się ona z pojęć, które – jako węzły w sieci – łączy struktura wzajemnych relacji i które mają także swoje powiązania z wyrazami tworzącymi kolejny poziom reprezentacji powstającej wypowiedzi. Dell (1986) unika dalszych spekulacji na temat tej reprezentacji, pojawiającej się na wczesnym etapie procesu mówienia, zakłada jednak, że budujące reprezentację semantyczną pojęcia muszą być zorganizowane w struktury tematyczno-rematyczne (mają przypisane określone role semantyczne), dzięki czemu na niższym poziomie możliwe jest przypisanie im konkretnych wyrazów pełniących określone funkcje w zdaniu (podmiot, orzeczenie, dopełnienie itd.).

Kolejny poziom, czyli kodowanie syntaktyczne (*syntactic encoding*) rozpoczyna się od wygenerowania ramy taktycznej odpowiadającej strukturze powierzchniowej zdania<sup>46</sup>. Rama taktyczna tego poziomu określa porządek wyrazów w zdaniu podyktowany zasadami gramatyki, tzn. tworzy uporządkowany linearnie zestaw „okienek” (*slots*) oznaczonych kategoriami gramatycznymi. Jednocześnie dokonywany jest wybór podstawowych jednostek leksykalnych, lemm, które wypełniają przygotowaną ramę taktyczną zgodnie z wyznaczonymi przez nią funkcjami w zdaniu. Selekcja lemm dokonywana jest na podstawie aktywacji pochodzącej od pojęć tworzących reprezentację semantyczną oraz ze względu na kategorie gramatyczne przypisane poszczególnym „okienkom” w ramie taktycznej tego poziomu. Pomyłki powstające na tym poziomie przygotowania wypowiedzi mogą polegać na aktywowaniu niewłaściwej lemmy (np. „pies” zamiast „kot”), nieodpowiednim wypełnieniu ramy taktycznej (np. zamiana miejscami wyrazów realizujących funkcję podmiotu i dopełnienia) lub błędnym skonstruowaniu ramy taktycznej zdania.

Dzięki temu, że lemmy wyposażone są w znaczniki (*markers*) kategorii gramatycznych możliwe jest przypisanie im odpowiednich morfemów, zgodnie z fleksją danego języka. Proces, w którym lemmy otrzymują swoją właściwą formę językową poprzez zastąpienie ich ciągiem morfemów odbywa się na kolejnym poziomie nazywanym przez Della (1986) kodowaniem morfologicznym (*morphological encoding*). Dla każdej aktywowanej i wybranej do realizacji podczas kodowania syntaktycznego lemmy tworzona jest rama taktyczna składająca się z „okienek”, wyznaczających miejsca dla morfemów rdzennych (podstawy słowotwórczej), a także przedrostków i przyrostków.

Dzięki temu, że podczas mówienia nadawca nie wybiera od razu całych leksemów ze słownika umysłowego, lecz dokonuje ich selekcji w dwóch krokach (najpierw dokonywana

---

46 Używając określenia „reprezentacja składniowa” (*syntactic representation*) Dell odnosi się do finalnego produktu procesów składniowych, który w terminologii gramatyki transformacyjno-generatywnej nazywany jest strukturą powierzchniową (Dell, 1986).

jest selekcja lemm a następnie przypisywane są im morfemy), możliwe jest twórcze i elastyczne posługiwanie się językiem nie tylko na poziomie różnorodności generowanych zdań, ale także – do pewnego stopnia – elastyczności w budowaniu planu wyrażania poprzez tworzenie neologizmów. Rodzi to jednak okazje do powstawania błędów skutkujących np. zamianą morfemów rdzennych lub formantów między sąsiadującymi wyrazami. Dell (1986) tłumaczy to niewłaściwym wypełnieniem ram taktycznych podczas kodowania morfologicznego, spowodowanym dysproporcją w poziomie aktywacji rozchodzącej się po sieci na skutek chwilowego zakłócenia np. spowodowanego zwiększonym tempem mówienia.

Po dokonaniu selekcji morfemów tworzona jest reprezentacja dźwiękowa wypowiedzi. Podczas kodowania fonologicznego (*phonological encoding*) zestaw morfemów zamieniany jest na ciągi fonemów umiejscawiane w ramy taktyczne składające się z „okienek” wyznaczających pozycje poszczególnych segmentów w strukturze sylaby. Kolejne aktywowane fonemy zajmują odpowiednio pozycje nagłosu (*onset*), ośrodka (*nucleus*) lub wygłosu (*coda*) w zależności od budowy konkretnej sylaby (sylaba otwarta vs zamknięta, sylaba lekka vs mocna). Traktując sylabę jako ramę taktyczną dla selekcji fonemów, a nie jednostkę kodowania, Dell (1986) wyjaśnia, dlaczego bardzo rzadko spotyka się w literaturze przykłady przejęzyczeń dokonywanych na sylabach. Według niego jednostkami kodowania są fonemy i to ich selekcja bywa czasem wadliwa (zamiana fonemów miejscami, antycypacje, perseweracje, substytucje segmentów). Sylaby natomiast stanowią struktury organizujące tę selekcję, a ich błędna konstrukcja podczas mówienia skutkuje przejęzyczeniami odbieranymi jako pomyłki w doborze fonemów.

Reprezentacja fonologiczna tworzy wzorzec dla instrukcji kierujących ruchem mięśni aparatu artykulacyjnego. Do rozpoczęcia artykulacji nie jest konieczne zakodowanie całej wypowiedzi na wyższych poziomach przetwarzania.. Dell (1986) nie wyjaśnia szczegółowo procesu artykulacji, najwięcej uwagi poświęcając kodowaniu fonologicznemu, którego przebieg został sformalizowany i przetestowany w szeregu symulacji porównywanych następnie z obserwacjami spontanicznych przejęzyczeń i wynikami badań laboratoryjnych.

W późniejszych latach Dell modyfikował swój model testując różne założenia na temat przebiegu procesu mówienia. Np. w modelu DoL (*division of labor*), w którym symulowane były procesy transformacji reprezentacji semantycznej na leksykalną (kodowanie syntaktyczne), zamiast pojęć stojących za poszczególnymi wyrazami, Dell i in. (2008) zastosowali sieć cech semantycznych, które dodatnio lub ujemnie wiązały się z poszczególnymi lemmami (np. cecha zdolności latania przynależy zarówno do wyrazu „ptak” i „samolot”, jednak cecha „zwierzę” pozytywnie wiąże się jedynie z wyrazem „ptak”

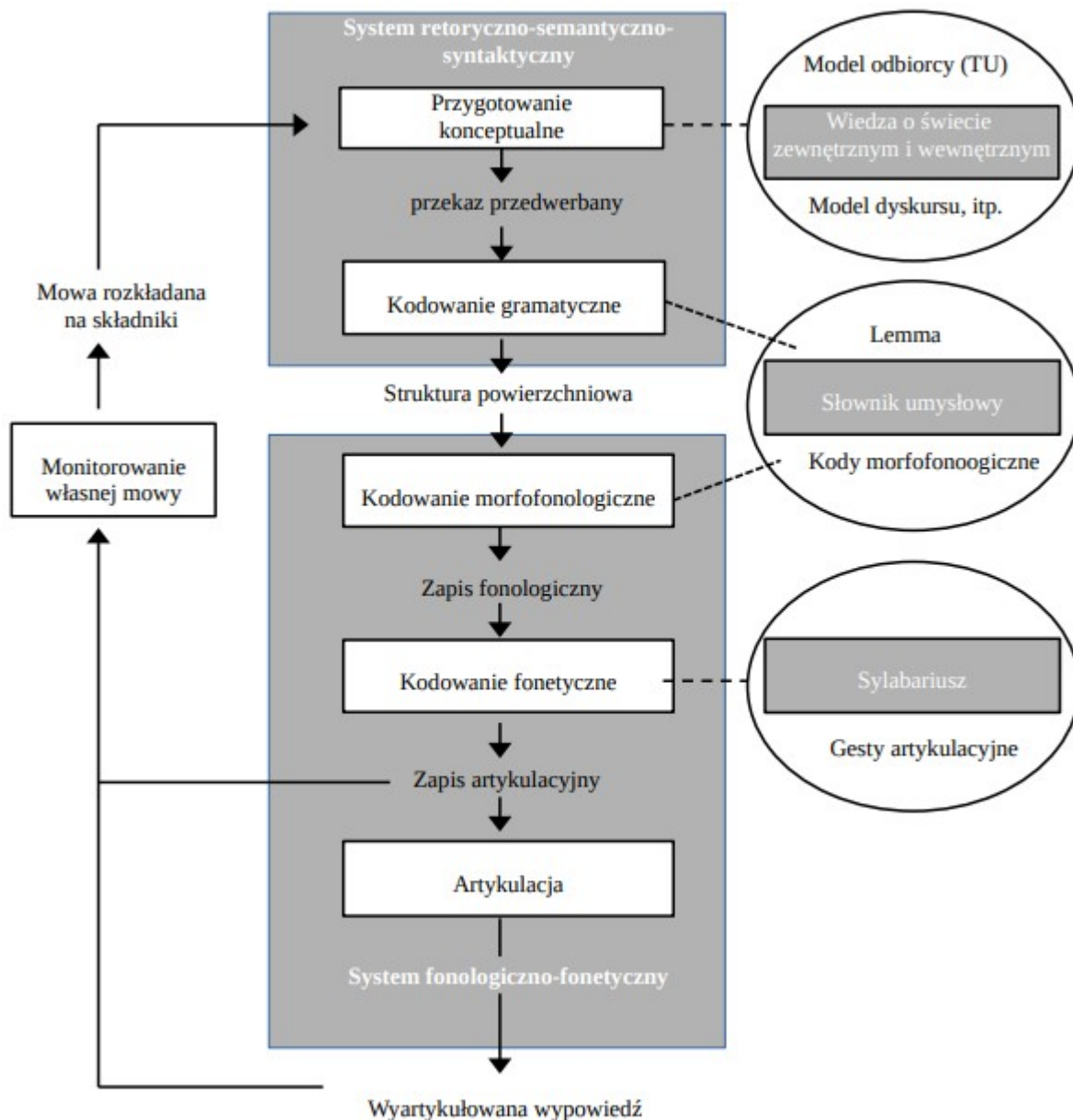
natomiast jej połączenie z wyrazem „samolot” ma w modelu wagę ujemną). Z kolei symulując błędy w nazywaniu obrazków przez afatyków i osoby zdrowe, Dell i in. (1997) całkowicie pominęli poziom kodowania morfologicznego zakładając bezpośrednie przejście z poziomu kodowania syntaktycznego na poziom kodowania fonologicznego (nazywanie obrazków polega na przygotowaniu do realizacji podstawowej formy leksemu i nie wymaga, aby została ona odmieniona przez przypadki, przez co reprezentacja morfologiczna składałaby się jedynie z morfemów rdzennych i jej pominięcie w modelu obliczeniowym jest możliwe bez straty dla jego trafności). Niezmienne natomiast pozostają podstawowe założenia o złożoności procesu mówienia (wielopoziomowość), interakcyjnym charakterze powiązań między poziomami, sieciowych relacjach łączących jednostki kodowania, aktywacji, która w sposób kaskadowy rozprzestrzenia się po sieci dokonując selekcji jednostek do ram taktycznych powstających na poszczególnych poziomach.

#### **2.2.4 Model produkcji mowy Levelta i in. (1999)**

Teoria mówienia zaprezentowana przez Levelta w 1989 roku była zwieńczeniem pewnego etapu pracy jej autora i podsumowaniem kilkudziesięciu lat badań z różnych obszarów psychologii, językoznawstwa i psycholingwistyki, które związane były z mową i przetwarzaniem języka. Zaprezentowany wtedy kompletny model produkcji mowy stał się przyczynkiem do dalszych badań i dyskusji teoretycznych, które ostatecznie doprowadziły do jego modyfikacji. Pierwszym istotnym krokiem było opracowanie w 1992 roku przez Roelofsa, doktoranta Levelta, obliczeniowego modelu uzyskiwania dostępu do lemm (Roelofs, 2005b). Model ten miał być istotnym wkładem w formalizację teorii Levelta (1989), jednak odchodził od niektórych pierwotnych założeń (np. odwoływał się do zjawiska rozprzestrzeniającej się aktywacji), w związku z czym zaczął funkcjonować jako niezależna teoria wydobywania lemm ze słownika umysłowego. Następnie Roelofs (2005b) opracował obliczeniowy model kodowania fonologicznego o nazwie WEAVER (akronim od słów *Word-form Encoding by Activation and VERification*). Z połączenia tych dwóch modeli symulujących proces leksykalizacji powstał ujednolicony model WEAVER++ i wraz z opisującą go ogólną teorią dostępu leksykalnego został opublikowany przez Levelta i współpracowników w 1999 roku (Levelt i in. 1999). W tym samym roku Levelt (1999b) opublikował projekt (*blueprint*) procesu mówienia uwzględniający zmiany jakie w stosunku do pierwotnego modelu z 1989 roku wprowadził model WEAVER++.

## Rysunek 4

Model produkcji mowy według Levelta (1999)



Źródło: Levelt (1999b), tłumaczenie za: Kurcz (2011)

Poniżej zaprezentowany zostanie opis procesu mówienia, jaki wyłania się z publikowanych przez Levelta i współpracowników prac na przełomie XX i XXI wieku.

Levelt (1999b; 2000) zrezygnował z nazw poszczególnych komponentów (modułów) na rzecz wyodrębnienia osobnych poziomów, które określane są przez pryzmat operacji, jakich dokonują. W nowej wersji modelu wyróżnia się kilka poziomów produkcji mowy,

które zagregowane są w dwa systemy: retoryczno-semantyczno-syntaktyczny<sup>47</sup> (pojęciowy) i fonologiczno-fonetyczny (językowy). Ten pierwszy odpowiedzialny jest za wydobycie z systemów wiedzy (np. pamięci) informacji (semantycznych i syntaktycznych), które posłużą do opracowania struktury powierzchniowej wypowiedzi. W tej postaci wypowiedź zostaje poddana procesowi kodowania językowego (otrzymuje językową formę), po którym następuje jej wyartykułowanie w postaci mowy zewnętrznej (*overt speech*).

Na pierwszym poziomie dokonuje się konceptualne przygotowanie wypowiedzi (*conceptual preparation*), czyli opracowanie przekazu przedwerbalnego. Na tym etapie nadawca decyduje o tym, co zamierza powiedzieć, tzn. jakie informacje chce przekazać (a jakich nie) oraz w jakiej kolejności (Levelt, 2001b). Selekcja informacji dokonuje się na podstawie różnych przesłanek związanych z wiedzą o zewnętrznym i wewnętrznym świecie, do których Levelt (1999b) zalicza rozumienie stanu umysłu rozmówcy (związane z posiadaną przez rozmówców teorią umysłu), przyjęty model dyskursu i inne sytuacyjne uwarunkowania pragmatyczne, skłaniające nadawcę do wyrażenia aktu mowy o określonej modalności i w określony sposób. Analogicznie do wcześniejszego modelu za ustalenie intencji komunikacyjnej, wygenerowanie celów i podcelów, selekcję informacji i ułożenie z nich przekazu o strukturze predykatowo-argumentowej odpowiada makro- i mikroplanowanie (Levelt, 1989; 1999b).

Podczas konceptualnego przygotowania wypowiedzi dobierane do przekazania informacje reprezentowane są w postaci aktywowanych pojęć leksykalnych (*lexical concepts*), którym w danym języku odpowiadają konkretne słowa. Model WEAVER++ zakłada, że poszczególne pojęcia leksykalne łączy sieć semantyczna, po której przemieszcza się aktywacja wzbudzając skojarzone ze sobą (sąsiadujące w sieci) pojęcia (Levelt i in., 1999). Pojęcia leksykalne występują jako węzły (*nodes*), które nie są reprezentowane przez zestawy cech semantycznych<sup>48</sup>, ale stanowią niepodzielne jednostki. Efektem etapu konceptualnego przygotowania jest przekaz przedwerbalny składający się z linearnie ułożonych pojęć leksykalnych (Levelt, 2001b).

Kolejnym krokiem generowania wypowiedzi językowej jest kodowanie gramatyczne (*grammatical encoding*). Wzbudzone pojęcia leksykalne aktywują odpowiadające im jednostki lemma, co stanowi pierwszy etap procesu selekcji leksykalnej (*lexical selection*).

---

47 Wszystkie elementy modelu zostaną zaprezentowane w tłumaczeniu na język polski za publikacją Kurcz (2011).

48 Niektóre teorie opisujące funkcjonowanie słownika umysłowego odwołują się do tzw. cech semantycznych (*semantic features*) lub semów, czyli podstawowych i nierozkładalnych na prostsze jednostek znaczenia, które tworzą znaczenia wyrazów (Kurcz, 1987; 2005). Choć dane pojęcie leksykalne można przedstawić jako zestaw cech semantycznych, na poziomie przekazu przedwerbalnego jest ono reprezentowane w tym modelu jako niepodzielna całość (węzeł sieci semantycznej).

Według Levelta i in., (1999) w proces selekcji leksykalnej wbudowany jest mechanizm weryfikacji, który kontroluje adekwatność aktywowanej lemmy względem realizowanego pojęcia leksykalnego. Lemma o najwyższym potencjale aktywacji może zostać wybrana do realizacji, kiedy pozwala na to mechanizm weryfikacyjny, co zabezpiecza system produkcji mowy przed błędami.

Podobnie, jak w modelu z 1989 roku leksykalizacja, czyli uzyskiwanie dostępu leksykalnego jest procesem dwuetapowym, składającym się z selekcji lemm i odpowiadających im form wyrazowych, jednak w nowszej propozycji postuluje się nieco inne właściwości lemm. We wcześniejszym modelu termin „lemma” odnosił się do semantycznych i syntaktycznych właściwości wyrazu, którego fonologiczną formę stanowił leksem. W nowszej wersji modelu Levelt i in. (1999) lemma, oznacza jedynie składnię wyrazu (*word's syntax*), a zatem informację o cechach składniowych wydobywanego ze słownika wyrazu (w przypadku czasowników jest to np. osoba, czas itd.). Za semantyczną stronę wyrazu w całości odpowiadają pojęcia leksykalne, a każdemu pojęciu przypisana jest konkretna lemma, która w procesie kodowania gramatycznego (*grammatical encoding*) zostaje ulokowana w odpowiednim miejscu tworzonej ramy syntaktycznej. Kodowanie gramatyczne przebiega linearnie, w sposób „przyrostowy” (*incrementally*)<sup>49</sup>.

Efektem kodowania gramatycznego jest struktura powierzchniowa (*surface structure*) wypowiedzi (Levelt, 1999b). Według Levelta i in. (1999) ten poziom reprezentacji nie musi jeszcze bezpośrednio wiązać się z mówieniem, a do jego aktywowania nie jest konieczna intencja komunikacyjna w znaczeniu powiedzenia czegoś na głos do innej osoby. Jest to specyficzny rodzaj myślenia w postaci „mówienia do siebie” (*speaking for thinking*), który można porównać do postulowanego przez Wygotskiego (1934/1989) „myślenia werbalnego”, czyli „mowy wewnętrznej”<sup>50</sup>. Warto w tym miejscu dodać, że dotychczas opisane poziomy produkcji mowy są współdzielone razem z systemem odbioru mowy. Inaczej mówiąc, sieci produkcji i percepcji mowy „pokrywają się” (są tożsame) od poziomu lemma „w górę”

---

49 Przyrostowy (*incrementally*) oznacza według Levelta (1999), że każdy kolejny komponent przetwarzania informacji może rozpocząć pracę na niekompletnym (nieukończonym) wyniku pracy poprzedniego procesu. Dany poziom produkcji mowy może rozpocząć przetwarzanie wypowiedzi mając na wstępie (na wejściu) dostępny jedynie jej fragment opracowany na wcześniejszym etapie, podczas gdy pozostała część wypowiedzi nie „opuściła” jeszcze poziomu wcześniejszego.

50 Termin „mowa wewnętrzna” jest używany przez Wygotskiego (1934/1989) w innym znaczeniu niż w pracy Levelta i in. (1999), którzy wykorzystują to określenie w odniesieniu do wyrazów fonologicznych, czyli reprezentacji powstałej w procesie kodowania fonologicznego. Pewnym pomostem między tymi dwiema koncepcjami może być propozycja Fernyhough, aby odróżnić mowę wewnętrzną rozszerzoną (*expanded inner speech*), zawierającą pełną informację językową, od skondensowanej mowy wewnętrznej (*condensed inner speech*), która pozbawiona jest cech fonologicznych, prozodii innych wskazówek artykulacyjnych (Geva i Fernyhough, 2019).

(Levelt, 1999b; Levelt i in., 1999), ale są rozdzielne dla kolejnych pięter wchodzących w skład systemu fonologiczno-fonetycznego.

Jak już wspomniano model Levelta (1999b; 2001b) postuluje dwuetapowy proces wyboru jednostek leksykalnych. W procesie kodowania formy wyrazu, czyli podczas kodowania fonologicznego (*phonological encoding*), następującego po kodowaniu gramatycznym, do aktywowanych lemm dobierane są odpowiadające im morfemy<sup>51</sup>. Proces ten ukierunkowany jest przez gramatyczne charakterystyki lemm, co pozwala na aktywację nie tylko morfemów leksykalnych, ale i fleksyjnych, czy słowotwórczych. Najpierw aktywowane morfemy zamieniane są na ciągi fonemów, którym przypisywane są cechy prozodyczne (kod fonologiczny oraz kod metryczny). Tu również model Levelta i in. (1999) przewiduje działanie mechanizmów weryfikacyjnych.

W kolejnym kroku odbywa się sylabifikacja (*syllabification*), czyli podział wygenerowanych ciągów fonemów na jednostki artykulacyjne niepokrywające się dokładnie z poszczególnymi wyrazami<sup>52</sup>, lecz z tworzącymi te wyrazy (w ich morfo-fonologicznym otoczeniu sąsiednich wyrazów) sylabami. Produktem końcowym kodowania fonologicznego są wyrazy fonologiczne (*phonological words*), czyli segmenty fonemów zawarte pomiędzy potencjalnymi pauzami (mogą być większe lub mniejsze od wyrazów leksykalnych), wyróżnione ze względu na swoją przynależność sylabiczną i przypisane im akcenty (tylko jedna sylaba w wyrazie fonologicznym jest nosicielem akcentu głównego; por. Polański, 2003). Z tych subleksykalnych segmentów budowane są w kodowaniu fonologicznym złożone frazy intonacyjne, które przekraczają granice poszczególnych wyrazów (Levelt, 2000). Kodowanie fonologiczne jest więc odpowiedzialne zarówno za tworzenie ramy fonologicznej wypowiedzi (wyrazy fonologiczne), służącej aktywowaniu odpowiednich sylab (proces sylabifikacji) jak i za nadanie jej cech suprasegmentalnych (proces prozodyfikacji), co po części uzależnione jest od modalności aktu mowy w ramach którego wypowiedź powstaje (np. intonacja wznosząca w pytaniach i opadająca w wypowiedziach twierdzących).

W pierwszym modelu Levelta (1989) kodowanie fonologiczne obejmowało również ustalanie cech fonetycznych generowanej wypowiedzi i bezpośrednio poprzedzało artykulację. W późniejszej propozycji teoretycznej z kodowania fonologicznego wydzielono

---

51 W przeciwieństwie do wcześniejszej koncepcji (Levelt, 1989) zdecydowano się zrezygnować z terminu „leksem”, a w jego miejsce, na oznaczenie formy wyrazu wydobywanej ze słownika umysłowego w drugim etapie dostępu leksykalnego, używać wyrazu „morfem”.

52 Podczas płynnej mowy artykułowane sylaby i cechy prozodyczne wypowiedzi przekraczają granice poszczególnych wyrazów. Przykładowo wyrazy „pisk opon” są wypowiedziane jako jeden wyraz fonologiczny złożony z trzech sylab: /pis/, /ko/, /pon/. Co więcej np. w krakowsko-poznańskiej odmianie polszczyzny słyszy się efekt udźwięcznienia w pierwszej sylabie, będącej wygłosem wyrazu poprzedzającego wyraz o nagłosie samogłoskowym: /piz/, /go/ /pon/. Efekt ten powstaje podczas kodowania fonetycznego, jako konsekwencja specyficznego doboru sylab.

kodowanie fonetyczne (*phonetic encoding*), które wedle Levelta i in. (1999) stanowi odrębny proces. Na tym etapie, kiedy ciągi fonemów mają już przypisaną strukturę sylabiczną i właściwości prozodyczne (strukturę metryczną), z magazynu nazywanego umysłowym sylabariuszem (*mental syllabary*) wydobywane są wzorce artykulacyjne poszczególnych sylab, z których składa się wypowiedź (Levelt i in., 1999). Tworzą one plan fonetyczny, który stanowi program pracy dla aparatu artykulacyjnego. W przypadku wyrazów rzadko używanych sylaby mogą być tworzone na bieżąco, zamiast być wydobywane z pamięci.

Według Levelta (1999b) każdy segment w planie fonologicznym wypowiedzi aktywuje podczas kodowania fonetycznego sylaby, które zawierają odpowiadającą temu segmentowi głoskę. W związku z tym, że sąsiadujące ze sobą segmenty aktywują niezależnie od siebie tę samą sylabę, suma aktywacji pozwala wydobyć tę sylabę spośród innych aktywowanych w sylabariuszu wzorców artykulacji<sup>53</sup>.

Ostatnim etapem produkcji mowy jest pobudzenie odpowiednich efektorów aparatu artykulacyjnego, skutkujące ruchem jego mięśni i powstaniem fali akustycznej niosącej zakodowaną informację językową (Levelt, 2000). Levelt (1999b) oraz Levelt i in. (1999) nie precyzują tego procesu ani nie rozwijają jego opisu, odsyłając do wcześniejszego modelu (Levelt, 1989).

Jak wspomniano na początku opisu tego modelu produkcji mowy, poziomy przetwarzania (generowania) wypowiedzi zagregowane są w dwa systemy: retoryczno-semantyczno-syntaktyczny (pojęciowy) i fonologiczno-fonetyczny (językowy). Ich rozdzielność ma swoje źródła w ontogenezie. Według Levelta i in. (1999) organizacja poziomów przetwarzania w procesie mówienia odzwierciedla dualizm, polegający na rozdzieleniu systemu pojęciowego i artykulacyjnego w pierwszych kilkunastu miesiącach życia dziecka. Ich zdaniem dziecko uczy się wiedzy o świecie w toku interakcji z otoczeniem, a kształtując stopniowo kategorie poznawcze, buduje fundament pod pojęcia leksykalne (*lexical concepts*), które w pierwszych miesiącach życia nie mają jeszcze swoich etykiet werbalnych. Osobnym torem przebiega rozwój języka, polegający z jednej strony na „dostrajaniu się” dziecka do odbioru mowy ojczystej, a z drugiej na generowaniu coraz bardziej złożonych, powtarzalnych sekwencji dźwięków, które stopniowo przypominając sylaby języka, staną się podstawą artykulacji pierwszych wyrazów. Połączenie się tych dwóch, osobno rozwijających się systemów (pojęciowego i artykulacyjnego) odbywa się w

---

53 Np. segmenty /t/, /e/ i /l/ w wyrazie „butelka” aktywują niezależnie od siebie różne sylaby, które zawierają w sobie odpowiadające im głoski [t], [e] i [l]. Wspólnie aktywują one także sylabę [tel], która może zostać wyartykułowana dzięki temu, że uzyskuje łączny poziom aktywacji przewyższający poziom aktywacji pozostałych potencjalnych sylab.

akcie mowy. Więcej informacji na temat rozwoju mowy w oparciu o model Levelta i in. (1999; Levelt, 1998) zostanie podane w rozdziale 4.

### 2.2.5 Porównanie poznawczych modeli produkcji mowy

Przedstawiony powyżej subiektywny wybór modeli produkcji mowy nie jest kompletny, gdyż w literaturze spotkać można jeszcze wiele autorskich ujęć opisujących przebieg tego procesu. Zaprezentowane modele stanowią jednak propozycje będące podstawą i punktem wyjścia dla innych autorów, zajmujących się kwestiami szczegółowymi, np. dostępem leksykalnym, charakterem kodowania fonologicznego itd. Zamieszczona poniżej tabela 1 zawiera zestawienie omawianych modeli pod względem uwzględnianych w nich poziomów produkowania mowy i subprocesów wchodzących w skład procesu mówienia.

**Tabela 1**

*Zestawienie porównywanych modeli produkcji mowy*

<b>Poziomy kodowania wg Frey'a</b>	<b>Model Garretta</b>	<b>Model Levelta</b>	<b>Model Levelta i in.</b>	<b>Model Della</b>
Kodowanie semantyczne	Poziom przekazu	(konceptualizator) Poziom przekazu	<i>Przygotowanie konceptualne</i> Pojecie leksykalne	Reprezentacja semantyczna
Kodowanie leksykalne	Poziom funkcjonalny	(formulator) Poziom lemma	<i>Selekcja leksykalna</i> Lemma	<i>Kodowanie syntaktyczne</i>
Kodowanie morfemowe	Poziom pozycyjny	(formulator) Poziom leksemu	<i>Kodowanie morfologiczne</i> <sup>54</sup> Morfem <i>Kodowanie fonologiczne</i> Wyraz fonologiczny	<i>Kodowanie morfologiczne</i>
Kodowanie fonologiczne	Poziom fonetyczny	Plan fonetyczny	<i>Kodowanie fonetyczne</i> Zapis artykulacyjny (Plan fonetyczny)	<i>Kodowanie fonologiczne</i>
Kodowanie motoryczne	Poziom artykulacyjny	(artykulator) Program motoryczny	<i>Artykulacja</i> Fala dźwiękowa	<i>Artykulacja</i>

*Adnotacja: Źródło – opracowanie własne;*

54 Levelt (1999b) używa określenia „kodowanie morfofonologiczne” (por. Rysunek 4) w odniesieniu do procesu, który Levelt i in. (1999) rozdzielają na dwa oddzielne subprocesy: kodowanie morfologiczne i kodowanie fonologiczne z sylabifikacją.

Według Harleya (2001) pomiędzy badaczami zajmującymi się produkcją mowy panuje względna zgoda, że proces ten składa się z trzech podstawowych etapów: konceptualizacji, formułowania i artykulacji (Harley, 2001). Poziomy te odpowiadają przetwarzanym względnie niezależnie od siebie składnikom semantycznym, syntaktycznym leksykalnym i subleksykalnym wypowiedzi. Większość badaczy jest także zgodna, że selekcja leksykalna przebiega na przynajmniej dwóch odrębnych poziomach kodowania (Nozari, 2018). Podstawowa kontrowersja między większością modeli dotyczy tego, czy traktują one produkcję (i rozumienie) mowy jako proces modułarny, czy nie (podejście modułarne vs koneksjonistyczne). Pociąga to za sobą dwa podstawowe pytania: (1) Czy przetwarzanie komunikatu odbywa się na kolejnych poziomach równocześnie, a informacja z jednego poziomu przechodzi kaskadowo do kolejnego zanim zostanie w pełni opracowana na wcześniejszym? Czy raczej programowanie wypowiedzi ma charakter sekwencyjny, a pomiędzy kolejnymi etapami „przesyłane są” tylko efekty pracy (pełna informacja) na danym poziomie? (2) Czy programowana informacja przetwarzana jest tylko w jednym kierunku (od intencji do artykulacji), czy możliwy jest także jej powrót do etapu wcześniejszego na zasadzie np. sprzężenia zwrotnego (Harley, 2001; Levelt, 1993; 1999a; Vigliocco i Hartsuiker, 2005)?

Współcześnie dyskutowane modele produkcji mowy odchodzą od podejścia modułowego na rzecz traktowania produkcji mowy jako aktywacji rozchodzącej się między węzłami sieci. Najważniejsze różnice dotyczą tego, czy jest to proces interaktywny (aktywacja rozchodzi się kaskadowo między węzłami i może przebiegać w obu kierunkach), czy dyskretny i ograniczony do jednego kierunku. Modele nie są też zgodne co do tego, czy selekcja leksykalna odbywa się przez konkurencję (*competition*) aktywowanych jednostek. Dane uzyskiwane eksperymentalnie nie pozwalają jeszcze na jednoznaczne rozstrzygnięcia (por. Dhooge i Hartsuiker, 2010), a w zależności od przyjętej interpretacji traktowane są jako poparcie dla jednego lub drugiego typu przebiegu programowania mowy (por. dyskusja Nozari, 2020 i Roelofs, 2020a,b).

Możliwe jest bowiem, że wykazywana w eksperymentach globalna niezależność poziomów przetwarzania informacji językowej w produkcji mowy nie odzwierciedla „lokalnych” dwukierunkowych koneksji między poziomami (Harley, 2001). Z kolei postulowana przez koneksjonistów interakcyjność poziomów prowadząca np. do mieszanych błędów nie musi być stałą własnością sieci, ale zakłóceniem w normalnie niekaskadowym (*non-cascading*) procesie rozchodzenia się aktywacji (Levelt, 1999a).

Ponadto, jak zauważają Hartsuiker i Notebaert (2010) chociaż współczesne obliczeniowe modele produkcji mowy dostarczają dokładnego opisu kolejnych etapów generowania wypowiedzi, a także przewidują możliwe wzorce błędów, to nadal słabo wyjaśniają przyczyny powstawania niepełności w mówieniu. Konieczne są więc nie tylko dalsze badania, ale także syntetyczne opracowania teoretyczne.

Niezależnie od przyszłych rozstrzygnięć, faktem jest, że przynajmniej czasami produkcja mowy ulega różnym zakłóceniom prowadzącym do przejęzyczeń, zawahań, gubienia wątku wypowiedzi, czy tzw. efektu końca języka. Najczęściej nie prowadzi to jednak do poważniejszych problemów w komunikacji, a błędy szybko są wykrywane i usuwane. Jest to możliwe dzięki kontroli procesu mówienia, która zostanie scharakteryzowana w kolejnym rozdziale.

### 3 Kontrola procesu mówienia

Mówienie jest czynnością intencjonalną, a realizacja intencji komunikacyjnej jest nadrzędnym celem tworzenia wypowiedzi kierowanej do innych osób. Wiedza o tym, jaki skutek przyniosły wypowiedziane słowa stanowi źródło informacji zwrotnej dla nadawcy (Postma, 2000) i umożliwia adekwatne formułowanie kolejnych komunikatów. Wymiana komunikatów między rozmówcami (także niewerbalnych) i ich wzajemna interpretacja jest więc ważnym aspektem kontroli tego, o czym i w jaki sposób mówią uczestnicy konwersacji. Odnosząc się do analizy aktów mowy (Austin, 1962/1993), można stwierdzić, że reakcja interlokutora może być wykorzystana przez nadawcę jako informacja o tym, czy jego wypowiedź została odebrana, zrozumiana zgodnie z intencją i czy został osiągnięty zamierzony skutek. Na tej podstawie nadawca komunikatu może wnioskować o powodzeniu lub niepowodzeniu czynności lokucyjnej, illokucyjnej oraz perlokucyjnej, a także podjąć kroki naprawcze, np. skorygować, rozszerzyć lub chociażby powtórzyć swoją wypowiedź.

Poleganie na tej zewnętrznej strategii w ocenie przebiegu procesu mówienia ma szczególne znaczenie w okresie nabywania języka. Rodzice, opiekunowie oraz starsze rodzeństwo zwracają uwagę na nieprawidłowości w mówieniu, korygując błędy i dostarczając wzorców prawidłowej realizacji. Z czasem rola innych osób w kontroli procesu mówienia zmniejsza się, gdyż każdy zyskuje zdolność do samodzielnego monitorowania swoich wypowiedzi (por. MacDonald i in., 2012). Ta umiejętność usprawnia komunikację, gdyż pozwala nadawcy wcześniej skorygować jakiejkolwiek nieprawidłowości, a odbiorcy skupić się na tym, w jaki sposób rozumie on słowa nadawcy, bez konieczności ciągłego dopytywania i poprawiania błędów. Przypomnijmy, że dzięki dyskretności języka (Hockett, 1968) nawet drobna zmiana w obrębie realizacji najbardziej podstawowych elementów języka jakimi są fonemy może zmienić całkowicie sens wypowiedzi (np. ubezdźwięczenie pierwszej głoski w wyrazie „bułka” skutkuje wypowiedzeniem słowa o zupełnie innym znaczeniu, tj. „półka”). Czekanie na reakcję odbiorcy, aby wykryć i poprawić taki błąd całkowicie dezorganizowałoby komunikację i oddalałoby osobę mówiącą od realizacji intencji. Być może dlatego obserwuje się wyraźną preferencję osób mówiących do dokonywania autonapraw (*self-repairs*) względem mniej pożądanых korekt inicjowanych przez interlokutorów (*other-initiated self-repairs*) (Forrester, 2008). Według Laakso (2010) preferencja ta pojawia się już bardzo wcześnie w rozwoju dziecka, bo już po ukończeniu drugiego roku życia autonaprawy zaczynają dominować nad naprawami motywowanymi z zewnątrz.

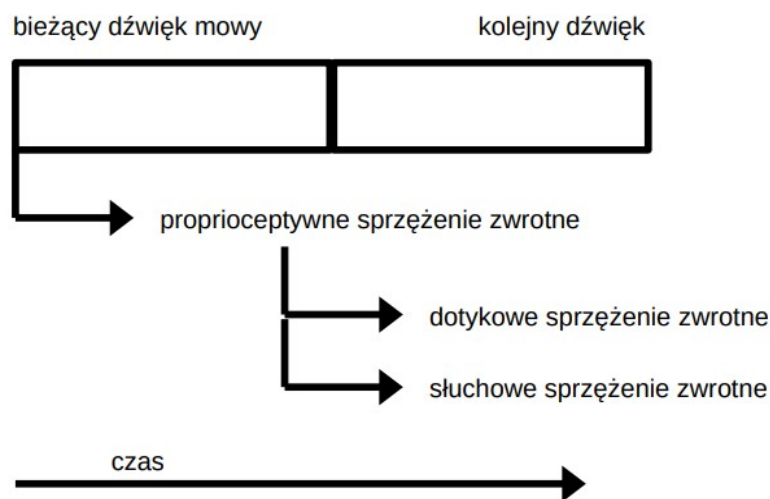
Jak zauważył Hockett (1979) ludzką komunikację cechuje całkowite sprzężenie zwrotne (*total feedback*), co oznacza, że osoba mówiąca sama ma możliwość odbioru generowanej przez siebie mowy. Z tego powodu mówienie, jak każda czynność intencjonalna może być kontrolowana przez osobę mówiącą – czyli wykonawcę czynności – samodzielnie, dzięki możliwości stałego monitorowania przebiegu tej czynności (Shao i in., 2012; Shao i in., 2013). Biorąc pod uwagę, że podstawowym kanałem mowy (wyłączywszy szczególne przypadki, np. komunikację osób głuchych) jest kanał głosowo-słuchowy (Hockett, 1979) informacja o wykonaniu czynności mówienia zwrotnie dociera do nadawcy właśnie tą drogą. W literaturze przedmiotu (Puppel, 1994; Postma, 2000; Lind i Hartsuiker, 2020) wyróżnia się trzy kanały zewnętrznej informacji zwrotnej (*external feedback*) składające się na kanał głosowo-słuchowy, które biorą udział w regulacji i kontroli przebiegu mówienia. Po pierwsze wrażenia zmysłowe związane z ruchem mięśni artykulacyjnych są rejestrowane przez kanał dotykowy (*tactile feedback channel*). Dostarcza on informacji o miejscu kontaktu struktur aparatu artykulacyjnego (np. zwarcia wargowego w artykulacji spółgłosek labialnych), sile i nacisku w kontakcie, a także o kierunku i kolejności ruchów artykulacyjnych. Kolejnym źródłem informacji są dane proprioceptywne (*proprioceptive feedback channel*) zbierane z receptorów mięśniowych (wrzecion mięśniowych) i stawowych, które rejestrują wielkość i szybkość rozciągania mięśni oraz kierunek i zakres ruchu w stawie skroniowo-żuchwowym. Trzecim źródłem informacji zwrotnej jest kanał słuchowego sprzężenia zwrotnego (*auditory feedback channel*), który odbiera dźwięki mowy na drodze przewodnictwa powietrznego i kostnego, rejestrując ich akustyczne parametry (częstotliwość, natężenie, periodyczność, czas trwania, czy kierunek docierania dźwięku).

Kanał proprioceptywny znajduje się pod kontrolą autonomicznego układu nerwowego (Postma, 2000) i reguluje pracę aparatu artykulacyjnego na bieżąco, zawiadując ruchem mięśni w trakcie generowania poszczególnych dźwięków mowy (Puppel, 1994). Kanały dotykowy i słuchowy odbierają informację już po wyartykułowaniu dźwięku, dlatego informacja zwrotna o wykonaniu jest rejestrowana i interpretowana już w trakcie artykulacji kolejnych głosek. Relacje czasowe w działaniu poszczególnych kanałów sprzężenia zwrotnego przedstawia rysunek 5.

Monitorowanie informacji płynących z trzech wspomnianych kanałów umożliwia prawidłową koordynację pracy aparatu artykulacyjnego w mówieniu z oddychaniem oraz dostrojenie artykulacji do wymogów otoczenia (np. regulacja głośności dźwięku w zależności od intensywności tła akustycznego). W przypadku, kiedy percepcja w którymś z kanałów

## Rysunek 5

### Kanały sprzężenia zwrotnego podczas mówienia



Źródło: opracowanie na podstawie Puppel (1994)

zostaje zakłócona w wyniku manipulacji eksperymentalnej (np. stosowanie obciążników modyfikujących ruch żuchwy lub podawanie słuchowej informacji zwrotnej z opóźnieniem, ang. *delay auditory feedback*) obserwuje się kompensujące działania aparatu artykulacyjnego (Schiller, 2005).

Choć kontrola aspektów prozodycznych mowy takich jak tempo, wysokość głosu oraz jego natężenie (głośność) mają niebagatelne znaczenie dla wywierania wpływu na odbiorcę komunikatu, to podstawowym nośnikiem treści jest tworzony przez nadawcę tekst wypowiedzi. Nadawca, aby więc ocenić, czy wypowiedziane słowa lub zdania realizują jego intencję, musi na bieżąco monitorować proces mówienia, oceniając czy to, co mówi jest zgodne z powziętym zamiarem oraz czy wypowiedź nie jest obciążona jakimś błędem (por. Levelt, 1989). Zdaniem Frydrychowicza (1999, s. 37) w momencie rozpoczęcia mówienia „mówiący wie tylko, o czym będzie mówił, ale nie wie jeszcze, co będzie mówił. Zna on tylko ogólny temat, przedmiot tego, o czym będzie mówił”. To ogólne pojęcie na temat treści wypowiedzi klaruje się wraz z nadawaniem jej postaci językowej, co stwarza okazję do zaistnienia sytuacji, w których nadawca usłyszy w swojej mowie nie to, co chciał powiedzieć. Odbiór i rozumienie własnej mowy bierze więc istotny udział w kontroli mówienia, gdyż

pozwała na porównanie efektu tego procesu ze stojącą u jego podstaw intencją. Co więcej, jak przekonują Lind i in. (2014), w warunkach eksperymentalnie wytworzonej, fałszywej informacji zwrotnej, część osób ufa bardziej temu, co słyszy w swojej mowie, niż temu, co rzeczywiście zostało przez nich powiedziane<sup>55</sup>.

### 3.1 Automonitorowanie procesu mówienia

Dokładna analiza błędów popełnianych podczas mówienia, która legła u podstaw poznawczych modeli produkcji mowy (zob. rozdział 2.), pozwoliła zauważyć, że spontaniczna mowa typowych użytkowników języka daleka jest od kompetencji „idealnego użytkownika języka” (Chomsky, 1985), tzn. bywa niegramatyczna, pełna pauz, zawahań i błędów, które – co ważne – często błyskawicznie są poprawiane (Postma, Oomen, 2005). Według Blackmer i Mitton (1991), które analizowały zapisy rozmów prowadzonych na antenie kanadyjskiego radia CBC, osoby wypowiadające się podczas audycji dokonywały różnego rodzaju korekt swoich wypowiedzi średnio co 5 sekund. Pokazuje to, że mowa spontaniczna nie tylko odbiega od standardów poprawności, ale jest na bieżąco oceniana i weryfikowana przez osobę mówiącą pod względem zgodności z tymi standardami. Innymi słowy, podczas mówienia nadawca w sposób ciągły i na bieżąco dokonuje inspekcji swoich wypowiedzi, a w razie wykrycia błędu lub niezgodności z planem dokonuje korekty (Hartsuiker i in., 2005; Lind i Hartsuiker, 2020). Co więcej, takie autoreperacje (*self-repairs*) są dokonywane bez zewnętrznych podpowiedzi ze strony innych rozmówców (np. w postaci zwrócenia uwagi na popełniony błąd), w bardzo krótkim czasie po wystąpieniu błędu (Postma, 2000).

Najogólniej można więc powiedzieć, że monitorowanie własnej mowy lub automonitorowanie<sup>56</sup> (*self-monitoring*) obejmuje proces lub zestaw procesów odpowiedzialnych za sprawdzanie poprawności generowanej przez nadawcę wypowiedzi oraz

---

55 W cytowanym badaniu osoby badane wykonywały test Stroopa, nazywając kolor napisu wyświetlającego się na ekranie komputera. W części prób zamiast swojej aktualnej wypowiedzi (np. *green*), słyszały podawane w słuchawkach nagrania (np. *grey*), które zostały wcześniej przygotowane z ich udziałem. W części przypadków osoby badane nie rejestrowały tego, że słyszany w słuchawkach ich własny głos przekazywał tekst o innej treści niż generowana przez nich mowa i akceptowały fałszywą słuchową informację zwrotną jako tożsamą z rzeczywistością wygenerowaną wypowiedzią (Lind i in., 2014).

56 Terminy „monitorowanie” oraz „automonitorowanie” będą stosowane wymiennie, jako synonimy na oznaczenie procesu monitorowania (przeglądania, sprawdzania, weryfikowania) tekstu własnej wypowiedzi. Ze względów teoretycznych ważne jest rozróżnienie na monitorowanie wewnętrzne i zewnętrzne (zob. dalej), a wyrażenie „automonitorowanie wewnętrzne” byłby tautologią, gdyż w przypadku monitorowania wewnętrznego można mówić wyłącznie o automonitorowaniu. Z tego względu jeśli w tekście nie zostanie zaznaczone, że „monitorowanie” odnosi się do odbioru i rozumienia mowy innych osób, termin ten będzie stosowany wyłącznie jako synonim automonitorowania.

pojęmowanie interwencji w razie niepowodzenia lub trudności w jej realizacji (Hartsuiker i Kolk, 2001; Lind i Hartsuiker, 2020). Monitorowanie służy temu, aby zweryfikować, czy do odbiorcy trafia tekst (produkt końcowy mówienia) zgodny z intencją nadawcy, a w razie wystąpienia w nim błędu (niezgodności z intencją), aby błąd został wykryty i poprawiony.

Dla wykrycia błędu kluczowe jest to, że powstała z jego udziałem wypowiedź nie spełnia w jakimś stopniu indywidualnych standardów poprawności przyjmowanych przez osobę mówiącą. Dwa podstawowe kryteria, które w tej ocenie są brane pod uwagę Vasić i Wijnen (2005) nazywają absolutnym i relatywnym. Kryterium absolutne odnosi się np. do poprawności gramatycznej wypowiedzi, natomiast kryterium relatywne dotyczy spójności komunikatu lub prawidłowej artykulacji. Oba kryteria mają charakter wewnętrzny, co znaczy że kluczowa jest nie tyle kwestia zgodności wypowiedzi z systemem językowym, ale ze standardami indywidualnymi osoby mówiącej. Część wykrywanych i korygowanych błędów mogłaby ująć uwagę odbiorców, gdyż nie łamią one ogólnie przyjętych reguł języka ani konwersacji, ale przez to, że wprowadzają jakiś rodzaj odchylenia od planu wypowiedzi, zniekształcają intencję nadawcy (relatywne kryterium popełnienia błędu). Patrząc z innej strony, błąd rozumiany jako złamanie jakiejś zasady językowej (np. błędna odmiana części mowy) może nie zostać zauważony ani poprawiony przez nadawcę, jeśli pozostaje w zgodzie z posiadaną przez niego wiedzą językową (nie łamie wewnętrznych standardów poprawności). Pierwszy rodzaj błędu będzie miał raczej charakter incydentalny, natomiast drugi będzie zazwyczaj błędem systematycznym, popełnianym notorycznie, a przez to zapewne niewykrywalnym dla monitorowania. Według Postmy (2000) istotne dla procesu monitorowania są błędy nie tylko niezamierzone, ale także w jakiś sposób nietypowe, wywołujące w nadawcy przeświadczenie, że coś „poszło nie tak” i że „mógłby, to zrobić lepiej”.

Dla wykrycia błędu może mieć znaczenie to, czy powstałe w trakcie programowania zakłócenie skutkuje wypowiedzeniem rzeczywistego słowa, czy zbitki głosek nie mających swojej reprezentacji w leksykonie umysłowym. Wielu autorów zajmujących się problemem produkcji i monitorowania mowy (np. Nootboom, 1980; Levelt, 1989; Postma i Oomen, 2005) zwraca uwagę na to, że wśród popełnianych przez użytkowników języka błędów fonologicznych występuje nadreprezentacja przejęzyczeń stanowiących rzeczywiste słowa (*lexical bias effect*) względem nic nieznaczających zbitok głosek (np. błędne powiedzenie „gałka” zamiast „grzałka” lub „niska” zamiast „miska”). Wśród różnych propozycji teoretycznych wyjaśniających to zjawisko (por. Dell, 1986) podnoszona jest także możliwość wstępnego filtrowania generowanej wypowiedzi pod względem jej „leksykalności”

(*lexicality*) (Hartsuiker i in., 2005). Błędy fonologiczne, które jednocześnie odpowiadałyby jakiemuś istniejącemu wyrazowi łatwiej „prześlizgiwałyby” się przez wewnętrzne systemy kontrolne, blokujące powstawanie błędów (omówienie ograniczeń w powstawaniu przejęzyczeń zob. Bręński, 2015).

Wykrycie błędu w wypowiedzianym tekście ma swoje konsekwencje dla toku mówienia, gdyż prowadzi do chwilowego przerwania artykulacji i wznowienia jej wedle programu realizującego korektę błędu. Według Vasić i Wijnena (2005) typowa korekta przejęzyczenia składa się z trzech faz: (a) interupcja mowy; (b) korekta błędu poprzez stworzenie nowego planu mowy; (c) restart artykulacji w miejscu zdania, w którym pojawiła się interupcja lub wcześniej. Dla zewnętrznego obserwatora dostępne są zatem trzy momenty świadczące o tym, że popełniony błąd został wykryty i poprawiony (Bręński, 2020):

1. popełnienie błędu – wyartykułowanie wypowiedzi (lub jej fragmentu) zawierającej błąd;
2. przerwanie artykulacji – zatrzymanie artykułowania kolejnych głosek po wykryciu błędu;
3. reperacja – wznowienie artykulacji wypowiedzi, w której zawarta jest korekta błędu.

Te trzy punkty zwrotne są ważnymi behawioralnymi wskaźnikami działania procesu monitorowania mowy. Przykładową wypowiedź z zaznaczonymi punktami zwrotnymi w monitorowaniu prezentuje rysunek 6. Jest to przykład wypowiedzi, w której osoba mówiąca opisuje swojemu rozmówcy kierunek, w którym ma się on udać. Zamiast powiedzieć „musi pan iść w lewo” błędnie stwierdza „musi pan iść w prawo” po czym przerywa mówienie (interupcja) i po chwili zawieszenia (hiatus) dokonuje reperacji błędu, mówiąc (dodając do wcześniej wypowiedzianych słów) „w lewo”. Jak zauważają Postma i Oomen (2005, s. 157) nadawca w takiej sytuacji „musi zrobić wiele rzeczy w ograniczonym czasie”, tj. musi zaplanować i rozpocząć wykonanie zamierzonej wypowiedzi, monitorować postęp w realizacji planu, wykryć błąd („w prawo”), zatrzymać artykulację, zreagować i zainicjować reperację, monitorować poprawność reperacji (i powtórzyć wcześniejsze kroki w przypadku gdyby reperacja była niepoprawna).

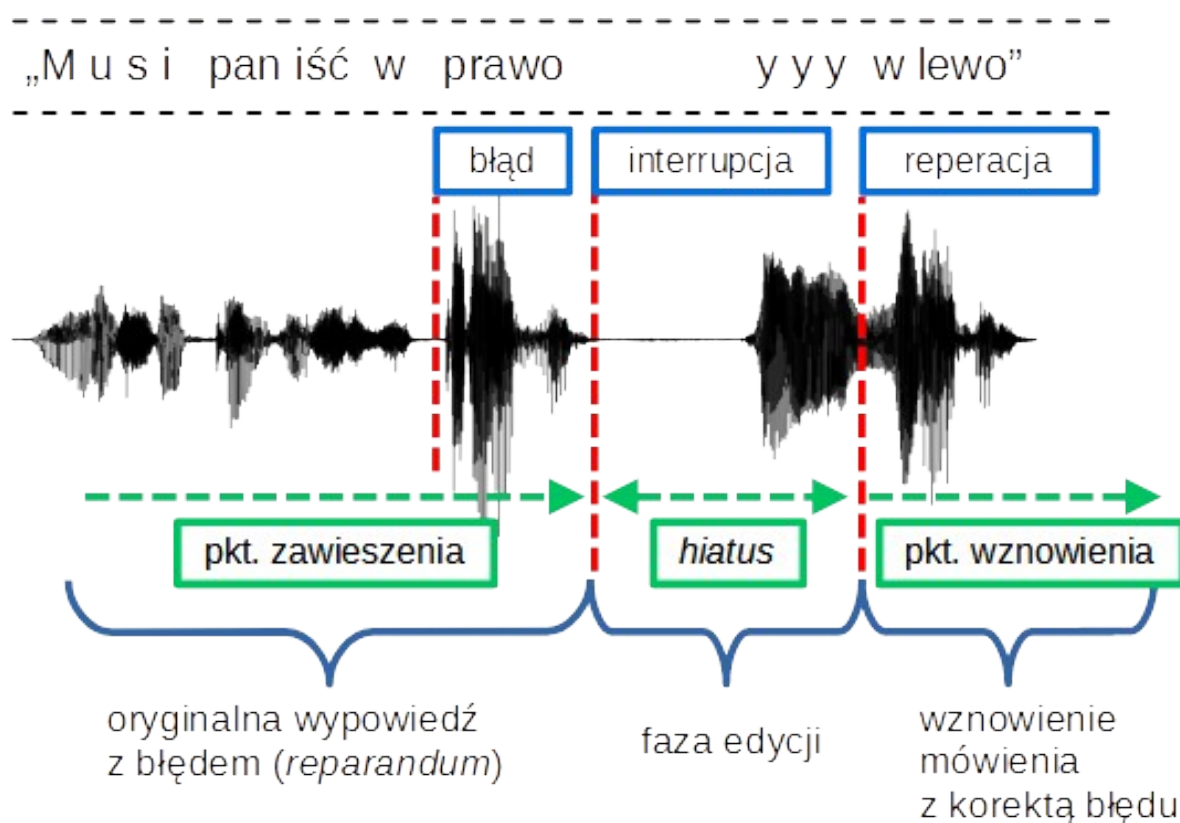
Na podstawie analizy spontanicznych reperacji przeprowadzonej przez Levelta (1983) można uznać, że przykład zaprezentowany powyżej wpisuje się w prototypową sekwencję: wypowiedź z błędem (*reparandum*) – interupcja – pauza – reperacja. Według Nootebooma (1980) natychmiastowe zatrzymanie toku mówienia tuż po wykryciu błędu jest pewną prawidłowością, którą nazywa główną zasadą przerywania mówienia (*Main Interruption Rule*): Zatrzymać strumień mowy natychmiast po wykryciu problemu<sup>57</sup> (Levelt, 1989, s. 478).

---

57 Org. *Stop the flow of speech immediately upon detecting trouble* (Levelt, 1989, s. 478)

## Rysunek 6

### Przykładowy błąd mowy z reperacją



Źródło: opracowanie własne na podstawie Bręński, (2020), Clark (1996), Levelt (1983), Vasić i Wijnen (2005)

W analizowanym przez Nootboom (1980) korpusie<sup>58</sup> większość błędnych wypowiedzi przerywana była tuż po wyartykułowaniu błędnego słowa. Analogicznie w badaniu Levelta (1983) w ponad 50% reperacji interrupcja następowała tuż po artykulacji słowa zawierającego błąd. Dodatkowo, zarówno Nootboom (1980), jak i Levelt (1983) odnotowali liczne przykłady<sup>59</sup>, w których strumień mowy przerywany był jeszcze przed pełnym wyartykułowaniem błędnego słowa.

Sekwencyjna analiza przedstawionego przykładu pozwala wyróżnić dwa istotne odcinki czasowe, które służą do wnioskowania na temat przebiegu ukrytych procesów monitorowania (por. Levelt, 1983; Hartsuiker i Kolk, 2001; Pillai, 2006): (1) czas pomiędzy wystąpieniem pomyłki, a interrupcją (*error-to-cutoff*) oraz (2) odcinek czasu pomiędzy

58 Nootboom (1980) analizował korpus opracowany przez Meringera i Meyera (zob. Cutler, 2006).

59 W analizowanym przez Nootboom (1980) korpusie przerwanie błędnego wyrazu jeszcze przed jego wypowiedzeniem dotyczyło 10% leksykalnych i 30% fonologicznych błędów mowy. Levelt (1983) odnotował zaprzestanie artykulacji przed pełnym wypowiedzeniem błędnego słowa w 18% przypadków.

przerwaniem artykulacji, a jej wznowieniem i realizacją reperacji (*cutoff-to-repair*), czyli czas trwania hiatusu. Zgodnie z główną zasadą przerywania mówienia (*MIR*) pierwszy odcinek czasowy wiązałby się z fazą detekcji błędu mowy (przerwanie wypowiedzi następuje natychmiast po wykryciu błędu), natomiast drugi odcinek, tzw. faza edycji (*editing phase*), wyznaczałby niezbędny czas potrzebny do zaprogramowania reperacji<sup>60</sup> (Levelt, 1983; Blackmer i Mitton, 1991).

Nooteboom (1980) oraz Levelt (1989) odnotowali jednak, że wypowiedzi nie są przerywane w dowolnym momencie (co sugerowałyby założenie o interupcji występującej natychmiast po wykryciu błędu), lecz zachowywana jest integralność wyrazów niezawierających błędu. Inaczej mówiąc, wypowiedź zawierająca błąd może być przzerwana w trakcie wypowiedzania błędnego wyrazu (w takiej sytuacji wypowiedź z rysunku 6. brzmiałaby np. *Musi pan iść w pra... yyy w lewo*), po tym wyrazie (jak na rysunku 6) lub po wypowiedzeniu w całości kolejnych wyrazów. Z kolei Blackmer i Mitton (1991), które analizowały korpus reperacji dokonywanych w naturalnej konwersacji, zauważyły że czas pomiędzy interupcją a reperacją może być praktycznie zerowy. Oznaczałoby to, że przerwanie wypowiedzi nie jest konieczne powiązane z detekcją błędu i nie występuje automatycznie po wykryciu pomyłki. Co więcej brak widocznej przerwy w mówieniu oznacza, że pauza nie jest konieczna dla zaplanowania reperacji i proces ten może się dokonywać w trakcie mówienia, tj. równoległe z realizacją wcześniejszego planu zanim jeszcze pojawi się interupcja (por. Postma, 2000; Hartsuiker i Kolk, 2001). Podważa to w istotny sposób uniwersalność głównej zasady przerywania mówienia (*MIR*).

Inna, ważna dla teorii monitorowania obserwacja pokazuje, że reperacje nie muszą dokonywać się wyłącznie zewnątrz (*overt repairs*). Już sam fakt, że wiele przejęzyczeń i pomyłek jest przerywanych w trakcie artykulacji błędnego wyrazu, każe zastanowić się, czy informacja o przebiegu mówienia dostarczana wspomnianymi wcześniej kanałami (dotykowym, proprioceptywnym, słuchowym) jest wystarczająca, aby wykryć błędy mowy tak wcześnie, tzn. zanim zostaną w pełni wyartykułowane. Ponadto obserwuje się takie wypowiedzi, które są przerywane i wznowiane, choć nie zawierają ewidentnego błędu. Levelt (1989) nazywa je ukrytymi reperacjami (*covert repairs*<sup>61</sup>), czyli reperacjami dokonywanymi na błędach, które zostały wykryte jeszcze przed ich wypowiedzeniem. W takich przypadkach artykulacja jest przerywana na podstawie antycypowanego błędu, zanim pojawi się on w

---

60 Wynika to z założenia, że zaplanowanie reperacji powinno odbyć się po wykryciu błędu (wcześniej nie ma podstaw aby dokonać korekty), ale przed wznowieniem artykulacji, będącej motoryczną realizacją nowego planu zawierającego korektę (Bręński, 2020).

61 W literaturze przedmiotu spotkać można także określenie „*prepairs*”, które nie jest jednak powszechnie używane (Hartsuiker, Kolk i Martensen, 2005) i nie ma ustalonego tłumaczenia na język polski.

słyszalnej mowie. To z kolei pociąga za sobą konieczność uznania, że monitorowanie może nie tylko operować zewnętrznie (*external monitoring*) na mowie transmitowanej kanałem głosowo-słuchowym, ale także działa wewnątrznie (*internal monitoring*) dokonując inspekcji przedartykulacyjnego planu wypowiedzi (Postma i Kolk, 1993). Istnienie mechanizmu wewnętrznego monitorowania procesu mówienia tłumaczyłoby także przypadki przerywania błędnych wyrazów w trakcie ich artykulacji, często zaledwie po wypowiedzeniu pierwszej głoski (Levelt, 1983; 1989)<sup>62</sup>. W takich sytuacjach błąd mógłby zostać wykryty jeszcze przed jego wypowiedzeniem na głos, jednak zbyt późno, aby powstrzymać rozpoczęcie artykulacji (Hartsuiker, Kolk i Martensen, 2005; Schiller, 2005). Kwestia tego, w jaki sposób działa monitorowanie wewnętrzne oraz z jakich źródeł informacji korzysta (jeśli nie opiera się na kanale głosowo-słuchowym) stało się punktem spornym między różnymi teoretycznymi modelami monitorowania, które zostaną opisane w następnym podrozdziale. W tym momencie opisane zostaną przesłanki świadczące o tym, że osoba mówiąca ma pewien dostęp do przedartykulacyjnego planu wypowiedzi i może kontrolować to, czy plan ten zostanie zrealizowany zewnętrznie (wyartykułowany).

Jak zauważa Postma (2000) ukryte reperacje są niejednoznaczne i trudne do klasyfikacji, gdyż w wypowiedzi nie uzewnętrznia się żaden błąd (brak jest reparandum), a tym samym trudno z całą pewnością mówić o reperacji tego błędu. Obserwowane zakłócenie mówienia w postaci interrupcji i restartu artykulacji jest poszlaką wskazującą na wykrycie w planie mowy błędu, jednak nie jest wystarczającym dowodem na to, że błąd rzeczywiście w planie wypowiedzi wystąpił (por. Engelhardt i in., 2013). Poza tym zakłócenia płynności mowy nie muszą być koniecznym przejawem wykrywania i korygowania antycypowanych błędów, ale mogą wynikać z trudności podczas planowania wypowiedzi na „wyższych piętrach” produkcji mowy (konceptualizacja, wybór odpowiednich słów itp.). Ponadto wykrywanie nadchodzących błędów programowania mowy i ich ukryte korygowanie może odbywać się na dwa sposoby: (1) poprzez strategię restartu (*restart strategy*), która przypomina zewnętrzne reperacje (wypowiedź jest przerywana i wznawiana po rewizji planu fonetycznego) lub (2) przez strategię odraczania (*postponement strategy*), która wiąże się z opóźnieniem artykulacji w celu rewizji przedartykulacyjnego planu mowy i zastąpienia błędnie wybranej jednostki, właściwą (Postma i Kolk, 1993). W takich przypadkach trudno

---

62 Levelt (1983, s. 64) podaje następujący przykład tego typu reperacji, który często (w tłumaczeniu na język angielski) przywoływany jest przez innych badaczy (np. Schiller, 2005; Hartsuiker, Kolk i Martensen, 2005):  
- „... zit een [ v ] . . . een horizontale lijn” (oryginalny przykład w języku niderlandzkim)  
- „... is a [ v ] . . . a horizontal line” (tłumaczenie na j. angielski)  
Zdaniem Levelta (1989) jest to przykład dla sytuacji, w której błędnie zaplanowany wyraz (*vertical* zamiast docelowego *horizontal*) zostaje przerwany zbyt szybko, aby jego wykrycie przypisać zewnętrznej pętli monitorowania.

jednoznacznie uznać, że spowolnienie tempa mówienia jest efektem działania wewnętrznego monitorowania, które chroni wypowiedź przed wprowadzeniem błędu do strumienia mowy. Z tego względu, aby uznać, że tzw. ukryte reperacje są rzeczywiście efektem działania monitorowania wewnętrznego, potrzebne są dowody jego działania w innych sytuacjach.

Inne dane na temat działania monitorowania wewnętrznego pochodzą z badań eksperymentalnych, w których osoby badane mają za zadanie wypowiadać jakiś tekst bezgłośnie („w myślach”) lub na głos, ale w warunkach maskowania odbioru słuchowego białym szumem (przegląd badań zob. Hartsuiker, Kolk i Martensen, 2005). W zadaniu, w którym głos uczestników jest maskowany są oni w stanie poprawiać pojawiające się błędy, pomimo tego że odbiór mowy drogą słuchową jest znacznie ograniczony<sup>63</sup>. Natomiast w zadaniu bezgłośnie mówienia, osoby badane są w stanie zgłaszać (poprzez naciśnięcie przycisku aparatury rejestrującej odpowiedzi) potencjalne błędy w realizacji planu mówienia, choć pojawiają się one wyłącznie w mowie wewnętrznej (bez artykulacji).

Zdolność do monitorowania wewnętrznego planu mowy, badana była również w kontekście innym niż wykrywanie błędów. Część aktualnej wiedzy o tym procesie pochodzi z badań nad przeglądaniem mowy wewnętrznej w poszukiwaniu wskazanych elementów subleksykalnych, np. sylab (Morgan i Wheeldon, 2003). Wheeldon i Levelt (1995) odkryli, że osoby badane są w stanie sprawdzać fonologiczną reprezentację generowanych bezgłośnie wyrazów, wskazując czy występują w niej podane segmenty.

Inny, często przywoływany przez badaczy (np. Levelt, 1989; Schiller, 2005; Roelofs, 2020) dowód działania monitorowania wewnętrznego pochodzi z badań Motleya i in., (1982), weryfikujących seksualne konotacje tzw. przejęzyczeń freudowskich (por. Motley, 1980). Badacze zastosowali procedurę SLIP, która polega na eksperymentalnym indukowaniu spuneryzmów. Procedura była zaprojektowana w taki sposób, aby możliwe było powstanie dwóch rodzajów przejęzyczeń: (1) wypowiedzi niestosownych (*obscene error utterances*, *taboo words*), czyli nacechowanych seksualnie lub wulgarnych oraz (2) neutralnych<sup>64</sup>. Osoby

---

63 Warto w tym miejscu odnotować, że droga akustyczna odbioru mowy obejmuje nie tylko przewodnictwo powietrzne (przenoszenie dźwięku do ucha wewnętrznego przez ucho środkowe i zewnętrzne), ale także przewodnictwo kostne (przenoszenie dźwięku do ucha środkowego w wyniku drgań mechanicznych czaszki wprowadzonych w rezonans). Maskowanie odbioru własnej mowy białym szumem ogranicza przewodnictwo powietrzne, jednak ma niewielki wpływ na przewodnictwo kostne, a poza tym nie blokuje informacji zwrotnej z kanałów dotykowego i proprioceptywnego. Można więc mieć uzasadnione wątpliwości, czy procedura maskowania białym szumem, rzeczywiście całkowicie ogranicza informację dostępną kanałem zewnętrznym (por. Hartsuiker, Kolk i Martensen, 2005) i wymaga od osoby mówiącej monitorowania wyłącznie wewnętrznego.

64 Procedura SLIP tworzy warunki eksperymentalne sprzyjające popełnianiu spuneryzmów (błędów zamiany głosek między wyrazami w parach, np. „mocny notyl” zamiast „nocny motyl”), powodując że osoba badana, kiedy ma np. wypowiedzieć parę słów „barn door” wypowiada słowa „darn bore”. Warunek neutralny w badaniu Motleya i in., (1982) sprzyjał popełnieniu błędów typu „cool tarts” (zamiast „tool carts”), natomiast w warunku primingu wyrazów niestosownych potencjalne błędy mogły skutkować wypowiedziami typu

badane, które brały udział w eksperymencie popełniały statystycznie mniej błędów niestosownych niż błędów neutralnych, gdyż istotnie częściej przerywały błędną wypowiedź nacechowaną seksualnie lub wulgarną niż wypowiedź neutralną, albo w ogóle nie dopuszczały do jej artykulacji. Co więcej, wskaźniki pobudzenia emocjonalnego (GSR) wskazywały istotnie większe pobudzenie u osób badanych, kiedy procedura skłaniała do popełnienia przejęzyczenia o charakterze tabu (nawet jeśli do przejęzyczenia nie dochodziło) niż wtedy, kiedy efektem manipulacji było wypowiedzenie pomyłki neutralnej. Dla autorów badania oraz innych badaczy zajmujących się tematem produkcji mowy i jej monitorowania (Levelt, 1989; Schiller, 2005; Roelofs, 2020), wyniki uzyskane przez Motleya i in., (1982) dowodzą, że wypowiedź może podlegać ocenie i redagowaniu (*editing*) zanim zostanie wyartykułowana. Zjawisko zwiększonego hamowania dla słów tabu, w porównaniu do neutralnych dystraktorów, wykazali także Severens i in. (2012) w badaniu wykorzystującym procedurę nazywania obrazków.

Badacze są więc zgodni, co do tego że osoba mówiąca ma dostęp do jakiejś formy planu przedartykulacyjnego, który może sprawdzić pod kątem poprawności językowej i zgodności z intencją. Odmienne są natomiast interpretacje tego z jakich źródeł informacji korzysta proces monitorowania wewnętrznego oraz na jak wczesnym etapie produkcji mowy może operować. Kwestia ta zostanie poruszona w następnym podrozdziale.

### **3.1.1 Modele monitorowania procesu mówienia**

Wśród badaczy zajmujących się tematem kontroli procesu mówienia panuje konsensus (Schiller, 2005; Nozari, 2020; Roelofs, 2020) dotyczący tego, że monitorowanie własnej mowy w oparciu o kanał słuchowy wiąże się z odbiorem i interpretacją informacji zwrotnej dokonywanymi przez system odbioru mowy (*speech comprehension system*), służący rejestracji wypowiedzi innych osób. Inaczej mówiąc nadawca ma możliwość słyszeć i rozumieć swoją wyartykułowaną wypowiedź, tak jak odbiera mowę innych osób (Huettig i Hartsuiker, 2010). Nie jest jednak jasne, jaka jest natura monitorowania wewnętrznego, które dokonuje inspekcji planu wypowiedzi jeszcze przed jej wyartykułowaniem.

Poznawcze teorie kontroli procesu mówienia używają pojęcia system edycji (redagowania) (*editing system*) oraz „monitor” (*monitor*) na określenie elementów architektury produkcji mowy, odpowiedzialnych za proces monitorowania. Np. Hartsuiker, Kolk i Lickley (2005, s. 277) stosują ten termin w odniesieniu do „dowolnego mechanizmu

---

„cool tits” (zamiast „tool kits”).

sygnalizującego rozbieżności między mową zamierzoną a wykonaną, niezależnie od tego, czy prowadzi on do pełnej analizy błędu<sup>65</sup>.

Blackmer i Mitton (1991) wyróżniają dwa typy monitorów, w zależności od sposobu ich działania. Monitor wstrzymujący (*a hold-up monitor*) może opóźniać dalsze przetwarzanie materiału podczas jego sprawdzania, a w razie wykrycia błędu anulować cały proces realizacji wypowiedzi. Monitor przepływowy (*a flow-through monitor*) może pozwolić na dalsze przetwarzanie materiału przy jednoczesnym jego sprawdzaniu. Uwaga badaczy zajmujących się spontanicznymi przejęzyczeniami i ich korygowaniem skupiona jest głównie na hamującej funkcji monitorowania, które powoduje przerwanie mówienia w efekcie wykrycia błędu, jednak w szerokim rozumieniu monitorowanie powinno mieć możliwość analizowania planu wypowiedzi, bez jej opóźniania lub wstrzymywania (por. Wheeldon i Levelt, 1995).

Centralnym problemem, przed którym stają poznawcze teorie procesu mówienia i jego kontroli jest to, aby nie traktować monitora jak *homunculusa*<sup>66</sup>, który odsłuchuje wypowiedź, śledząc ją w poszukiwaniu błędów oraz decydując jakie działanie podjąć (Postma i Kolk, 1993). Z tego względu poszczególne modele kontroli mowy proponują „zaawansowane monitory, które służą dość ograniczonym celom” (Postma, 2000, s. 124). Poznawcze teorie kontroli procesu mówienia różnią się pod względem liczby postulowanych monitorów, sposobu ich działania oraz źródła informacji zwrotnej o sposobie realizacji planu mówienia, z której monitorowanie korzysta (Hartsuiker i in., 2005). Inaczej mówiąc teorie te odmiennie odpowiadają na pytanie skąd osoba mówiąca dowiaduje się, że jej wypowiedź (jeszcze niewyartykułowana) zawiera błąd i czy rzeczywiście musi być świadoma tego błędu, aby zatrzymać wypowiedź i dokonać reperacji, czy też proces ten przebiega do pewnego stopnia automatycznie, czasem metodą prób i błędów (przykłady niedokładnych reperacji, zatrzymywanych i wznawianych aż do osiągnięcia akceptowalnej formy korekty), przybliżając wypowiedź do zamierzonego celu (Postma i Oomen, 2005). Analizując literaturę przedmiotu Postma (2000) wyróżnił trzy podstawowe podejścia do zagadnienia monitorowania mowy, które zostaną przedstawione w kolejnych akapitach tego rozdziału.

---

65 Org. We use the term 'monitor' here to denote any mechanism that signals discrepancies between intended and executed speech, irrespective of whether it produces a full analysis of the error (Hartsuiker, Kolk i Lickley, 2005, s. 277).

66 Homunculus – łac. *mały człowieczek* – typowy dla podejścia poznawczego problem wyjaśniania procesów samokontroli przez przypisywanie źródła tej kontroli jakiemuś złożonemu, nadrzędnemu konstruktowi, który de facto nie ma wartości eksplanacyjnej, a jedynie zastępuje wcześniejsze mało precyzyjne, intuicyjne pojęcia jak „wola”, czy „świadomość” (Chuderski, 2010).

### 3.1.1.1 Modele monitorowania oparte na produkcji mowy

Pierwsze, chronologicznie najwcześniejsze podejście zakłada, że monitorowanie jest wewnętrznym mechanizmem produkcji mowy (*production-based approach*), niezależnym od procesu odbioru mowy. Mechanizm ten ma bezpośredni dostęp do każdego poziomu przetwarzania mowy, dzięki czemu w momencie wykrycia błędu może naprawić wypowiedź rozpoczynając jej programowanie od etapu, na którym wystąpiło zakłócenie. Propozycja ta jest typowa dla wczesnych modeli produkcji mowy z lat 70-tych i 80-tych ubiegłego wieku (modele te nie były omawiane w rozdziale drugim).

Jeden z wariantów zakłada istnienie centralnego monitora lub edytora (*editor*), który sprawuje pieczę nad przebiegiem procesu mówienia. Miałby on mieć dostęp do każdego z poziomów programowania mowy i w razie wykrycia błędu, podejmować działania regulacyjne. Trudno jednak obronić taki element umysłu przed zarzutem o to, że jest to niewiele wyjaśniający przykład homunkulusa (Levelt, 1989). Ponadto, wczesne modele produkcji miały charakter modułarny, co zakładało enkapsulację informacji przetwarzanej w modułach i czyniło ją potencjalnie niedostępną dla innych modułów lub procesorów (Postma i Oomen, 2005). Trudno zatem teoretycznie pogodzić hermetyczność modułu z możliwością zewnętrznej ingerencji monitora w jego pracę, chyba że ograniczy się działanie monitora do sprawdzania informacji wyjściowej na każdym z etapów produkcji mowy. Np. model Kempena (van Wijk i Kempen, 1987) zakładał, że centralny monitor nadzoruje przepływ informacji między modułami, sprawdzając czy informacja wyjściowa z danego modułu nie zawiera błędu (por. Blackmer i Mitton, 1991).

Innym rozwiązaniem tego problemu było postulowanie istnienia nie jednego, centralnego monitora, mającego dostęp do każdego z poziomów programowania mowy, ale wielu autonomicznych i wyspecjalizowanych monitorów działających w obrębie poszczególnych pięter tego procesu (Postma i Oomen, 2005). W tej propozycji każdy z monitorów wykonuje swoją kontrolną pracę, ograniczoną do danego poziomu produkcji mowy. W razie wykrycia błędu, planowanie reperacji rozpoczyna się na poziomie, na którym błąd wystąpił. Proces korygowania błędów nie wymaga świadomości, gdyż odbywa się automatycznie. Podstawowy zarzut wobec tak sformułowanego mechanizmu monitorowania dotyczy jego niskiej ekonomiczności. Założenie, że na każdym z poziomów produkcji mowy działa niezależny monitor, który sprawdza i kontroluje przebieg kodowania informacji na tym poziomie, pociąga za sobą problem reduplikacji wiedzy (Postma, 2000). Na przykład, jeśli podczas pracy formulatora wybierana jest jednostka leksykalna ze słownika umysłowego, to

odpowiedzialny za kontrolę tego procesu monitor także musiałby mieć dostęp do słownika umysłowego, aby porównać przygotowany program z wzorcem (por. Levelt, 1989).

Krytyka z jaką spotkały się wczesne modele monitorowania oparte na produkcji (Postma, 2000; Postma i Oomen, 2005), spowodowała, że nie były dalej rozwijane i współcześnie mają one wartość głównie historyczną. W bieżących dyskusjach na temat monitorowania procesu mówienia określenie modele oparte na produkcji (*production-based*) odnosi się do modeli koneksjonistycznych, które zostaną krótko scharakteryzowane poniżej (Nozari, 2020; Roelofs, 2020a).

### **3.1.1.2 Koneksjonistyczne modele monitorowania produkcji mowy**

Kolejnym z głównych podejść do monitorowania wyróżnionych przez Postmę (2000) są teorie koneksjonistyczne bazujące na modelach produkcji mowy takich jak propozycja Della (1986, por. rozdział 2.). Podobnie jak wspomniane wyżej modele monitorowania opartego na produkcji mowy, teorie koneksjonistyczne także zakładają, że monitorowanie wewnętrzne jest umocowane w systemie produkcji mowy, jednak w przeciwieństwie do tych pierwszych, nie zakładają istnienia odrębnych monitorów, a mechanizmy kontroli (np. wykrywanie błędów i ich korygowanie) wywodzą wprost ze sposobu w jaki aktywacja rozprzestrzenia się po sieci leksykalnej. System ten jest samosterowny, tzn. jest w stanie regulować wszelkie zakłócenia (np. pomyłki) dzięki tym samym sprzężeniom zwrotnym, jakie zachodzą między poszczególnymi poziomami architektury produkcji mowy (Levelt, 1989).

Przykładowo, proces w wyniku którego dochodzi do pomyłki oraz jej poprawienia można opisać następująco: Po dokonaniu selekcji danego węzła w sieci (np. wybór lemmy) na bazie poziomu jego aktywacji, pewna porcja aktywacji jest przesyłana do hierarchicznie podporządkowanych mu węzłów (np. poziom kodowania morfologicznego). W wyniku tego procesu aktywowane są morfemy realizujące dany wyraz, a następnie (na niższym poziomie hierarchii) aktywowane są fonemy, które wchodzi w skład danego morfemu. Według Della (1986) błędy w procesie mówienia powstają w wyniku wzbudzenia niewłaściwego węzła. Zgodnie z założeniem o pozytywnym sprzężeniu zwrotnym między poziomami produkcji mowy, możliwe jest porównanie poziomu aktywacji przesłanej do węzłów podrzędnych z informacją otrzymywaną zwrotnie. Im większa rozbieżność między informacją o poziomie przesłanej niżej aktywacji, a zwrotną informacją o poziomie aktywacji wyselekcjonowanych węzłów podrzędnych, tym większa szansa na wykrycie błędu (Postma, 2000). Według Postmy

i Kolka (1993) zaletą takiego podejścia jest redukcja wykrywania błędów do prostego porównywania współczynników aktywacji, choć – podobnie jak w modelach modułarnych – podejście to zakłada mało ekonomiczne sprawdzanie każdego poziomu reprezentacji tworzonej wypowiedzi.

Zbliżony mechanizm, oparty na monitorowaniu poziomu aktywacji opisuje MacKay (1992). W swoim koneksjonistycznym modelu produkcji i percepcji mowy (węzły sieci są wspólne dla obu systemów) rozróżnia on terminy aktywacja (*activation*) oraz torowanie (*priming*). Ten pierwszy termin oznacza sytuację wzbudzenia węzła na zasadzie „wszystko albo nic”, przy czym aktywacja nie rozprzestrzenia się między węzłami, a jej intensywność nigdy się nie zmienia. Okres aktywacji węzła kończy się krótkim okresem zmniejszonej pobudliwości, którą MacKay nazywa autohamowaniem (*self-inhibition*). Tym, co wiąże węzły ze sobą, analogicznie do rozprzestrzeniającej się aktywacji w teorii Della (1986), jest torowanie. Torowanie ma zmienny poziom natężenia i jest niezbędne do tego, aby aktywować dany węzeł, który bez uzyskania odpowiedniego poziomu torowania pozostawałby nadal w stanie spoczynku. Aktywowany węzeł najsilniej toruje aktywację bezpośrednio z nim połączonych węzłów, które z kolei z mniejszą siłą torują wzbudzenie kolejnych połączonych z nimi węzłów itd. Torowanie z różnych źródeł ulega sumowaniu, więc o aktywacji danego węzła decyduje suma torowania płynącego z różnych powiązanych z nim węzłów. W tak zorganizowanej architekturze selekcja jednostek produkcji mowy odbywa się na zasadzie „najsilniej torowany wygrywa” (*the most-primed-wins*), co czyni ją podatną na występowanie błędów (MacKay, 1992, s. 50). Na przykład błąd substytucji może wystąpić na skutek tego, że odpowiedni z punktu widzenia realizacji intencji węzeł na danym poziomie nie otrzyma torowania wystarczającego do aktywacji i „przegra” z silniej torowanym, konkurencyjnym węzłem. Według MacKaya (1992) wykrycie błędu jest możliwe dzięki informacji zwrotnej, płynącej w górę hierarchicznej sieci z podrzędnych aktywowanych węzłów. Droga torowania i aktywowania węzłów na poszczególnych poziomach tworzy pewien szlak aktywacji, po którym powinna także przebiec informacja zwrotna. Kiedy więc na którymś z poziomów aktywowany zostanie niewłaściwy węzeł, przekazuje on pewną porcję torowania do wyższych węzłów, które pierwotnie nie były aktywowane. Tworzy to stan rozbieżności, gdyż dzięki mechanizmowi autohamowania, wcześniej aktywowane węzły powinny zostać wytłumione, aby zapobiec wielokrotnemu wybieraniu tego samego elementu. Kiedy system osiągnie w ten sposób nietypowy lub nieoczekiwany stan, jest to sygnałem dla osoby mówiącej, że w planie mowy pojawił się błąd (Nozari, 2018). Wykrywanie błędów opiera się na świadomości (*awareness*) przedłużonej aktywności węzłów, które pierwotnie nie były

zaangażowane w plan wypowiedzi, a które zwrótnie zostały aktywowane przez przepływ torowania z błędnie wybranych węzłów na niższych piętrach produkcji mowy (Postma, 2000). Przy czym u osoby mówiącej świadomość wystąpienia potencjalnego błędu jest wtórna wobec jego wykrycia, tzn. jest konsekwencją detekcji błędu przez system, a nie jej prekursorem (Nozari, 2018). Propozycja MacKaya (1992) ogranicza proces monitorowania wyłącznie do oceny poziomów aktywacji w ramach systemu produkcji mowy, bez uwzględnienia informacji płynących np. z zewnętrznej pętli mowy. Nie jest więc w stanie wyjaśnić różnic w monitorowaniu mowy spowodowanych modyfikacjami informacji o własnej wypowiedzi odbieranej kanałem słuchowym (Postma, 2000).

Współczesną kontynuacją i modyfikacją modeli Della i MacKaya jest teoria monitorowania konfliktu Nozari i in. (2011; Nozari, 2018, Nozari, 2020). Według Nozari (2018) monitorowanie jest procesem podejmowania decyzji o prawdopodobieństwie popełnienia błędu przy użyciu informacji generowanych w produkcji mowy. Detekcja błędów zależy więc od wewnętrznej dynamiki systemu produkcji mowy. Nozari podobnie jak MacKay uważa, że system, który rutynowo postępuje według znanych wzorców jest w stanie rozpoznać nieznaną wzorce, jako odstające od normy (Nozari, 2020). Nie musi przy tym kierować się kryterium prawidłowości generowanej wypowiedzi (nie porównuje efektu z zamierzeniami), a raczej szacuje prawdopodobieństwo błędu na bazie wielkości konfliktu, jaki pojawia się między wzorcami aktywacji (Nozari, 2018). Konflikt oznacza pewien stan systemu produkcji mowy, w którym aktywowane są alternatywne, konkurencyjne wobec siebie reprezentacje na różnych poziomach hierarchicznej sieci węzłów (np. leksykalnym, fonologicznym, składniowym itp.). W typowej sytuacji, kiedy wypowiedź jest tworzona bezproblemowo, jeden sposób reprezentacji na danym poziomie przetwarzania (dany węzeł w sieci) osiąga wyższą aktywność niż inne, mogące potencjalnie z nim konkurować. W takiej sytuacji z systemu płynie informacja o niskim poziomie konfliktu. Kiedy jednak wraz z aktywowanym węzłem (np. reprezentacja na poziomie jednostki leksykalnej) silnie wzbudzone są również inne węzły, dzielące z nim cechy semantyczne lub fonologiczne, występuje między tymi węzłami silniejszy konflikt i wzrasta ryzyko pomyłki (Hanley i in., 2016). Inaczej mówiąc, kiedy różne reprezentacje osiągają podobny poziom aktywacji, jest to sygnał, że podczas programowania wypowiedzi mogą pojawić się problemy, gdyż efekt końcowy procesu przestaje być jasny (Nozari, 2020). Konflikt wyrastający z naturalnej dynamiki systemu produkcji mowy jest więc sygnałem o potencjalnym ryzyku błędu, przy czym „system nie musi wiedzieć” jaka reakcja jest prawidłowa, aby informacja o konflikcie była użyteczna (Nozari, 2018). Im wyższy poziom generowanego przez system konfliktu

(mniejsza różnica w poziomie aktywacji konkurencyjnych węzłów), tym większe prawdopodobieństwo odczytania go przez centrum wykonawcze (przypuszczalnie zlokalizowane w korze przedniego zakrętu obręczy, *anterior cingulate cortex* – ACC) jako sygnał błędu, skłaniający osobę mówiącą do zatrzymania i rewizji swojej wypowiedzi (Hanley i in., 2016). Decyzja o tym jak silny konflikt jest wystarczająco wysoki, aby zostać wykrytym jako potencjalny błąd jest zadaniem dla ram decyzyjnych (*decision making framework*), które działają na bazie szacowanego prawdopodobieństwa wystąpienia błędu, bez konieczności angażowania świadomej uwagi (Nozari, 2020). W zależności od celów osoby mówiącej, kryteria oceniające poziom tolerowanego konfliktu mogą być różne. W sytuacji, kiedy liczy się mowa wolna od błędów kryteria mogą być bardzo restrykcyjne i z łatwością wykrywać każdy potencjalny błąd, co jednak będzie zwiększać także prawdopodobieństwo wystąpienia fałszywych alarmów (Nozari, 2018). Z kolei tolerowanie małej różnicy w poziomie aktywacji, a więc dopuszczanie do występowania konfliktu rodzi ryzyko wystąpienia jawnych przejęzyczeń. Te z kolei wykrywane są już przez zewnętrzny monitor związany z odbiorem mowy.

Jak zauważa Roelofs (2020a) propozycja Nozari zakłada zatem istnienie dwóch mechanizmów monitorowania: pierwszy, podstawowy mechanizm monitorowania wewnętrznego oparty na wykrywaniu konfliktu w systemie produkcji mowy oraz drugi, bazujący na odbiorze mowy podczas monitorowania zrealizowanych wypowiedzi oraz mowy wewnętrznej wypowiedzianej bezgłośnie. Konsekwencją przyjmowanych przez Nozari i in. (2011) założeń jest sytuacja, w której osoba nie wie czy sygnał alarmu rzeczywiście zwiastował błąd, dopóki faktycznie nie przejęzyczy się i nie usłyszy tej pomyłki.

### **3.1.1.3 Teoria pętli percepcyjnej – monitorowanie oparte na systemie odbioru mowy**

Trzecie i ostatnie podejście wyodrębnione przez Postmę (2000) tym różni się od poprzednich, że cały proces monitorowania wiąże z systemem odbioru i rozumienia mowy (*speech comprehension system*). Najszerzej dyskutowanym współcześnie modelem opartym na systemie odbioru mowy (por. Hartsuiker i Kolk, 2001; Nozari i in., 2011; Postma, 2000) jest teoria pętli percepcyjnej (*Perceptual Loop Theory*), zaproponowana przez Levelta (1983; 1989) i rozwijana później przez Hartsuikera (2007; Hartsuiker i Kolk, 2001; Hartsuiker, Kolk i Martensen, 2005) oraz Roelofsa w ramach modelu WEAVER++ (Levelt i in., 1999; Roelofs, 2005; 2020a).

Postulowana przez Levelta pętla percepcyjna ściśle powiązana jest z jego modelem procesu mówienia, przedstawionym w rozdziale 2 (na rys. 2.2). W przeciwieństwie jednak, do wcześniej zaprezentowanych teorii monitorowania opartych na procesie produkcji mowy, propozycja Levelta (1983; 1999) bazuje na monitorze korzystającym z informacji dostarczanych przez system odbioru i rozumienia mowy (*comprehension-based monitoring*). Takie rozwiązanie było odpowiedzią na problemy modeli wielomonitorowych, w których – jak wcześniej wskazywano – wiedza ulegałaby reduplikacji. Według Levelta (1989) osoba mówiąca może rejestrować swoją własną mowę tak samo, jak rejestruje mowę innych osób. Aby ocenić aspekty semantyczne, składniowe, fonologiczne itd. mowy wewnętrznej (przed artykulacją), nie są potrzebne dodatkowe monitory operujące na poszczególnych piętrach procesu produkcji mowy, lecz możliwe jest wykorzystanie tych samych rozwiązań, które służą rozumieniu mowy. System odbioru mowy umożliwia nie tylko odczytanie treści wiadomości ze strumienia mowy, ale także jest wrażliwy na odchylenia od językowych standardów. (Levelt, 1989). Najkrótszą formułę działania pętli percepcyjnej Levelt (1983, s. 50) wyraża w zdaniu, że „wewnętrzna i zewnętrzna mowa jest postrzegana, analizowana i sprawdzana pod kątem adekwatności zamierzonej i kontekstowej, zgodności z zamierzonym i dostarczonym przekazem oraz poprawności językowej”<sup>67</sup>. Centralny monitor kontrolujący przebieg procesu mówienia zlokalizowany jest w konceptualizatorze. Oznacza to, że konceptualizator, który inicjuje generowanie wypowiedzi sprawdza także poprawność postępu jej realizacji (Postma i Kolk, 1993). W przeciwieństwie do formulatora i jego komponentów, które w wysoce zautomatyzowany sposób odpowiedzialne są za kreowanie językowej reprezentacji wypowiedzi, konceptualizator znajduje się pod częściową świadomą kontrolą (Vasić i Wijnen, 2005), co czyni monitorowanie także procesem (częściowo) świadomym i powiązaniem z centralną kontrolą (Mattson i Baars, 1992).

W modelu Levelta (1983; 1989) są więc trzy ścieżki lub kanały wejściowe do monitora.

1. Pierwsza ścieżka związana jest z bezpośrednim dostępem do efektów pracy konceptualizatora, pozwalająca sprawdzić przedwerbalny przekaz pod względem jego odpowiedniości względem intencji, wiedzy nadawcy oraz kontekstu. Przejawem zewnętrznym działania monitora w tym zakresie są reperacje dostosowujące (*appropriateness repairs*)<sup>68</sup>. Dwie pozostałe ścieżki należą do systemu odbioru mowy – stąd Levelt (1983) używa także określenia „podwójnej pętli percepcyjnej” – i służą one ewaluowaniu

67 Org. In conclusion, then, we prefer to assume that repairing speech involves a perceptual loop: the self-produced inner or overt speech is perceived, parsed and checked with respect to intentional and contextual appropriateness, agreement of intended and delivered message, and linguistic correctness (Levelt, 1983, s. 50).

wypowiedzi posiadającej formę językową otrzymaną w procesie formułowania. Pętla wewnętrzna (*internal loop, inner loop*) obejmuje przedartykulacyjny plan fonetyczny<sup>69</sup>, czyli mowę wewnętrzną. Na rysunku jest naszkicowana jako połączenie pomiędzy buforem artykulacyjnym, który okresowo przechowuje plan fonetyczny, w czasie gdy czeka na artykulację, a systemem rozumienia mowy (por. Hartsuiker i in., 2005). Pętla zewnętrzna (*external loop / outer loop*) obejmuje mowę generowaną przy użyciu artykulatora, którą mogą usłyszeć i zrozumieć za pomocą własnego systemu rozumienia mowy wszyscy ludzie będący w zasięgu głosu osoby mówiącej, łącznie z nią samą. Rolą systemu rozumienia mowy jest dokonanie analizy planu fonetycznego mowy wewnętrznej (etap przedartykulacyjny procesu mówienia) oraz ciągów fonetycznych, odebranych przez analizator słuchowy (etap postartykulacyjny procesu mówienia). Dzięki tej analizie, obie pętle dostarczają do monitora mowę rozczłonkowaną (*parsed*), która może zostać porównana z zamierzoną wypowiedzią na poziomie konceptualizacji (Postma i Kolk, 1993; Hartsuiker i in., 2005, por. Roelofs, 2020a). Jak zauważa Hartsuiker (2007) teoria Levelta nie precyzuje szczegółowo tego, jak przebiega proces porównywania (por. Roelofs, 2020a).

Zaletą tego rozwiązania jest to, że bazuje na wspólnym dla mowy zewnętrznej i wewnętrznej mechanizmie odbioru mowy podczas monitorowania i nie postuluje żadnych dodatkowych monitorów, które miałyby czuwać nad przebiegiem procesu mówienia na różnych jego etapach. Jedynie końcowy (przedartykulacyjny) plan fonetyczny (mowa wewnętrzna) oraz mowa zewnętrzna mogą być przetwarzane przez system rozumienia mowy, ten sam, który służy do przetwarzania mowy odbieranej z zewnątrz (Levelt, 1989; Postma, 2000).

Levelt zakłada, że pętla dla mowy zewnętrznej jest dłuższa niż dla wewnętrznej (przesłanie danych do monitora wymaga więcej czasu), gdyż angażuje się w wyodrębnianie ze strumienia dźwięku cech akustycznych oraz artykulacyjnych mowy, do których z kolei nie ma dostępu pętla wewnętrzna. Z tego względu pętla dla mowy wewnętrznej działa szybciej (np. wcześniej jest w stanie wykryć błędy), ale jednocześnie jest mniej dokładna pod względem wykrywania błędów fonetycznych (np. błędne udźwięcznienie lub ubezdźwięcznienie spółgłosek albo perturbacje w realizacji samogłosek) (Levelt, 1989).

---

68 Ten typ reperacji nie dotyczy popełnienia błędu, ale jest przejawem doprecyzowywania przez nadawcę treści, która została wprowadzona przez niego do wypowiedzi w sposób nieadekwatny z punktu widzenia odbiorcy (np. wprowadzała dwuznaczność).

69 W pierwszej wersji swojej teorii Levelt (1989) zakładał, że monitorowanie wewnętrzne ma dostęp do planu fonetycznego. W dalszych badaniach (Wheeldon i Levelt, 1995) pojawiły się przesłanki, że pętla wewnętrzna może analizować także bardziej abstrakcyjną reprezentację tworzoną przez wyrazy fonologiczne. Zostało to uwzględnione w kolejnej wersji teorii (Levelt i in., 1999; Levelt, 1999b; por. Postma, 2000; Schiller, 2005), która zostanie przedstawiona w dalszej części tego podrozdziału (por. rozdział 2).

Biorąc pod uwagę fakt, że pętla zewnętrzna jest w stanie wykryć jakiegokolwiek błąd dopiero po ich wyartykułowaniu oraz założenie o tym, że pętla wewnętrzna operuje na produkcie końcowym procesu kodowania językowego (plan fonetyczny), dostępnego tuż przed artykulacją, istotne staje się pytanie o efektywność pętli wewnętrznej we wczesnym wykrywaniu błędów. Rozwiązanie zaproponowane przez Levelta zakłada bardzo krótkie okienko czasowe dla pracy pętli wewnętrznej, równoważne z czasem buforowania planu fonetycznego przed artykulacją. Jak zauważa Postma (2000) w teorii percepcyjnej, wykrycie błędu w odniesieniu do momentu jego zewnętrznego wystąpienia zależy od dostępności oraz pojemności bufora artykulacyjnego, a zwiększenie tempa mówienia zmniejsza zakres buforowania. Kwestia czasu jaki pętla wewnętrzna ma do dyspozycji zależy więc od tego, jak bardzo planowanie fonetyczne wyprzedza artykulację. Według Levelta pojemność bufora artykulacyjnego wynosi od jednej do kilku fraz fonologicznych, a minimalny czas potrzebny Artykulatorowi na „rozpakowanie” programu fonetycznego i rozpoczęcie artykulacji, to 200-250 ms. Jest to bardzo krótki czas pozwalający na ewaluację wewnętrznego planu mowy, ale niekoniecznie na podjęcie działań zapobiegających przejęzyczeniu. Jak wspomniano wcześniej, część wypowiedzi zawierających błąd może być przerywana w czasie artykulacji błędnego wyrazu lub tuż po nim. Levelt (1983; 1989) interpretuje te przypadki, jako efekt wczesnego wykrycia nieprawidłowości jeszcze przez pętlę wewnętrzną (przed artykulacją), ale na tyle późno, że program artykulacyjny już został uruchomiony i możliwe jest dokonanie jedynie reperacji zewnętrznej. Podczas szybkiego mówienia, może nie być w ogóle czasu na buforowanie planu fonetycznego, stąd impuls do interupcji może pojawić się już po rozpoczęciu artykulacji. Z kolei przy niskim tempie mówienia interwał pomiędzy zakończeniem formułowania planu fonetycznego, a rozpoczęciem artykulacji będzie wystarczający, aby wykryć i poprawić błędy przed ich ujawnieniem w mowie zewnętrznej (Blackmer i Mitton, 1991).

W ten sposób osoba mówiąca może monitorować wszystkie istotne parametry wypowiedzi od treści, przez gramatykę, dobór słów oraz ich właściwą wymowę. Według Levelta dostęp do tych wszystkich aspektów wypowiedzi nie jest ciągły i jednoczesny. Osoba mówiąca może słyszeć swoje artykułowane słowa (tak, jak słyszą je inni) lub monitorować plan mowy wewnętrznej, tuż przed artykulacją. Na tej podstawie, korzystając z systemu rozumienia mowy może odczytać wszystkie istotne parametry, służące ewaluacji wypowiedzi. Ostateczną instancją oceniającą wypowiedź jest intencja komunikacyjna, dlatego niektóre błędy lub pomyłki mogą zostać pominięte (czasem także nie są rozpoznane). Z tego też względu monitor jest powiązany z konceptualizatorem i mając bezpośredni dostęp do treści

przedweralnego przekazu może korygować planowane dopiero wiadomości zanim rozpocznie się ich językowe formułowanie.

Pierwotnie, teoria Levelta (1983; 1989) przyjmowała działanie opisanej przez Nootebooma (1980) głównej zasady przerywania mówienia (*MIR*). Badania Blackmer i Mitton (1991) dostarczyły jednak danych podważających uniwersalność tej reguły. Badaczki zarejestrowały wiele przykładów pomyłek i ich reperacji, które nie wpisywały się w kanoniczny schemat „reparandum – interupcja – pauza – reperacja” (por. Levelt, 1983). Po pierwsze, wbrew wspomnianej zasadzie, część przejęzyczeń nie jest konieczne zatrzymywanych w trakcie ich artykulacji lub tuż po ich wypowiedzeniu, ale obserwowalna jest tendencja do odraczania interupcji. Potwierdzają to obserwacje Seyfeddinipur i Kita (2001), które analizowały gestykulację towarzyszącą mówieniu. Według nich struktura i dynamika gestów ulega zakłóceniu już chwilę przed wstrzymaniem artykulacji, co potwierdza, że wykrycie błędu nie musi oznaczać natychmiastowej interupcji. Po drugie, w wielu przypadkach czas pomiędzy zatrzymaniem mówienia, a jego wznowieniem (*cutoff-to-repair*) bywa znikomy, zbyt krótki, aby tę przerwę efektywnie wykorzystać do zaplanowania korekty. Według Blackmer i Mitton (1991) planowanie korekty rozpoczyna się jeszcze przed wstrzymaniem artykulacji (a nie występuje dopiero po nim), o czym świadczą liczne przypadki reperacji występujących natychmiast po interupcji, bez widocznej pauzy lub opóźnienia. Na tej podstawie Postma i Kolk (1993) sformułowali alternatywną zasadę opóźnionej interupcji (*delayed-interruption principle*), która zakłada rozpoczęcie planowania korekty po wykryciu błędu w planie wewnętrznym mówienia, ale jeszcze przed dokonaniem interupcji (co może być czynnikiem tę interupcję opóźniającym). Regułę tę, jako zmodyfikowaną główną zasadę przerywania mówienia (*modified main interruption rule*) Hartsuiker i Kolk (2001; Hartsuiker, 2007) włączyli do swojego obliczeniowego modelu pętli percepcyjnej (*perception loop model*), stanowiącego formalizację teorii Levelta (1983; 1989). Przyjęli oni założenie, że planowanie reperacji oraz przygotowywanie interupcji stanowią dwa równoległe procesy uruchamiane niezwłocznie wraz z detekcją błędu. Hartsuiker i Kolk podjęli się także oszacowania wkładu, jaki w monitorowanie i korygowanie błędów mają pętla wewnętrzna i zewnętrzna. Chociaż pierwotnie przyjęli, że monitorowanie wewnętrzne i zewnętrzne jest tak samo skuteczne, w późniejszej pracy (Hartsuiker, Kolk i Martensen, 2005), dopasowując opracowany model do danych empirycznych, doszli do wniosku, że istnieją różnice w podziale pracy (*division of labor*) między pętlami. Według nich monitorowanie mowy wewnętrznej wymaga większego wysiłku (*effort*), gdyż służy głównie prewencyjnemu zapobieganiu wystąpienia błędów. Monitorowanie zewnętrzne może jedynie

wykryć błąd już po jego wypowiedzeniu, zatem nie jest w stanie uchronić płynności wypowiedzi przed niepożądanymi zakłóceniami (Hartsuiker, 2007). Jednocześnie zewnętrzna pętla percepcyjna ma dostęp do danych fonetyczno-akustycznych (por. Levelt, 1989), do których nie ma dostępu pętla wewnętrzna, dlatego skuteczniej może wykrywać błędy fonologiczne, których nie udało się zidentyfikować przed artykulacją. Z tego względu Hartsuiker, Kolk i Martensen (2005) wnioskują, że generalnie wewnętrzny kanał monitorowania konsumuje więcej zasobów, dzięki czemu wykrywa więcej błędów (szczególnie leksykalnych) niż kanał zewnętrzny. Wyjątkiem są błędy fonologiczne, które są trudniej rozpoznawane w pętli wewnętrznej – w ich przypadku obie pętle działają z równą skutecznością (por. Roelofs, 2020a).

Inne zmiany w teorii pętli percepcyjnej wprowadzone zostały wraz z przedstawieniem uaktualnionej wersji teorii produkcji mowy (teoria Levelta powiązana z modelem kodowania fonologicznego WEAVER ++, opisana w rozdziale 2.) autorstwa Levelta i in. (1999). Podstawowa różnica między obiema wersjami modelu dotyczy reprezentacji, do której dostęp ma monitorowanie wewnętrzne. We wcześniejszej propozycji, Levelt (1989) twierdził, że pętla monitorowania wewnętrznego ma dostęp do mowy wewnętrznej, reprezentowanej przez plan fonetyczny, dostarczany do bufora artykulacyjnego przez formulator. Badanie Wheeldon i Levelta (1995) pokazało jednak, że reprezentacja, na której operuje monitorowanie ma charakter bardziej abstrakcyjnego wyrazu fonologicznego. Osoby badane podczas wykonywania zadania eksperymentalnego osiągały dostęp do poszczególnych fonemów w generowanych bezgłośnie słowach (poprzez tłumaczenie słyszanych słów angielskich na język niderlandzki) rozpoznając je w sposób przyrostowy (*incremental*), co odzwierciedla sposób w jaki postępuje kodowanie fonologiczne (sukcesywne umieszczanie segmentów fonologicznych w „ramie” wyrazu). Jednocześnie wykonanie zadania nie ulegało pogorszeniu w trakcie głośnego odliczania, angażującego artykulację (a więc i plan fonetyczny mowy)<sup>70</sup>, wrażliwe natomiast było na granice sylabowe, co znaczy, że materiał poddany monitorowaniu wewnętrznemu musiał przejść proces prozodyfikacji. W teorii Levelta i in. (1999) poziomem reprezentacji, który spełnia te warunki jest wyraz fonologiczny (por. rozdział 2.), a interpretacja wyników badania Wheeldon i Levelta (1995) zgodna jest z późniejszymi badaniami nad monitorowaniem elementów subleksykalnych w generowanych wypowiedziach (Morgan i Wheeldon, 2003; Schiller, 2005). Według Levelta (1999) pętla

---

70 Wheeldon i Levelt (1995) dowiedli, że pomimo zadania dodatkowego, jakim było głośne odliczanie, efekt przyrostowego dostępu do poszczególnych fonemów w generowanych wyrazach został zachowany. Jednocześnie odnotowali opóźnienie w wykonaniu tego zadania, spowodowane dodatkowym zadaniem odliczania, co sugeruje, że zasoby z których korzysta monitorowanie mogą być ograniczone przez inne zadania. Problem ten będzie dyskutowany w kolejnym podrozdziale.

monitorowania wewnętrznego ma więc dostęp do wyrazów fonologicznych, będących efektem kodowania morfo-fonologicznego, poprzedzającego kodowanie fonetyczne, służące przygotowaniu instrukcji dla bufora artykulacyjnego.

Pozostałe różnice w opisie działania pętli percepcyjnej w dwóch kolejnych odsłonach teorii Levelta (1989; 1999), wynikają przede wszystkim ze zmian w rozumieniu architektury poznawczej konstytuującej proces mówienia. Włączenie modelu WEAVER++ do opisu produkcji mowy pozwala modelować przebieg tego procesu (przynajmniej w jego części), jak i odbioru mowy, na zasadzie rozprzestrzeniającej się aktywacji między węzłami sieci reprezentującymi poszczególne jednostki programowania mowy. To do pewnego stopnia otwiera możliwość istnienia wzajemnych powiązań między systemami produkcji i rozumienia mowy, postulowanych wyraźnie w modelach koneksjonistycznych, gdyż kolejne piętra przetwarzania języka nie są traktowane tak hermetycznie, jak w typowo modułarnych teoriach. Jak zauważa Postma (2000) zaproponowane modyfikacje umożliwiają większą interakcję między produkcją i percepcją mowy, niż to zakładał model pętli percepcyjnej w pierwotnej postaci (Levelt, 1983; 1989; por. Roelofs, 2020a). W nowszej wersji teorii postulowane jest współdzielenie przez system produkcji i odbioru mowy części systemu retoryczno-semantyczno-syntaktycznego, to znaczy sieci produkcji i percepcji pokrywają się od poziomu lemma „w górę”<sup>71</sup> (Levelt i in. 1999). Poza tym nadal utrzymywana jest odrębność produkcji i odbioru mowy na poziomie systemu fonologiczno-fonetycznego, choć jednocześnie dopuszcza się istnienie ograniczonych powiązań między nimi (Levelt, 1999b). Z tego względu niezmiennie pozostaje także założenie, że monitorowanie własnej mowy odbywa się za pośrednictwem systemu odbioru mowy, który mając do dyspozycji dane słuchowe (zewnętrzna pętla percepcyjna) oraz dane z mowy wewnętrznej reprezentowanej przez wyrazy fonologiczne (wewnętrzna pętla percepcyjna), może dokonać ewaluacji wypowiedzi pod względem zgodności z intencją i wewnętrznymi standardami poprawności językowej.

### **3.1.2 Porównanie modeli monitorowania procesu mówienia**

Przedstawione modele monitorowania różnią się pod względem przyjmowanych założeń na temat tego, w jaki sposób osoba mówiąca radzi sobie z problemem powstawania błędów mowy i ich naprawianiem. Najszerzej akceptowanym współcześnie modelem monitorowania procesu mówienia jest teoria pętli percepcyjnej (por. Postma, 2000; Hartsuiker i Kolk, 2001; Nozari i in., 2011). Z kolei największa krytyka spotkała pierwsze modele opisujące proces

---

<sup>71</sup> Org. (...) all production lemmas are perceptual lemmas; the perceptual and production networks coincide from the lemma level upwards (Levelt i in., 1999, s. 7).

monitorowania jako edytowanie wypowiedzi na różnych poziomach produkcji mowy. Zarzuty dotyczyły przede wszystkim tego, że postulowanie mechanizmu opartego na wielu monitorach godzi w zasadę ekonomii poznawczej (np. problem reduplikacji) i jest trudne do empirycznego zweryfikowania (Levelt, 1989; Postma, 2000). Poza tym, skoro efekty pracy poszczególnych modułów, czy procesorów generujących wypowiedź miałyby być porównywane z jakimś wzorcem, oznaczałoby to, że w systemie istnieje już prawidłowa reprezentacja kodowanego elementu, zatem powinna być ona wykorzystana jako pierwsza, co skutecznie wyeliminowałoby jakiegokolwiek błędy (MacKay, 1992; por. Postma i Kolk, 1993). W związku z tym, że te wczesne rozwiązania teoretyczne, traktujące monitorowanie jako wewnętrzny mechanizm kontrolny procesu mówienia, mają współcześnie wartość jedynie historyczną, nie będą dalej dyskutowane.

Pozostałe podejścia lepiej poradziły sobie z krytyką, choć jak przyznaje Levelt (1992) w odpowiedzi na zarzuty MacKaya, zarówno jego własna teoria pętli percepcyjnej, jak i rozwiązania oferowane przez teorie koneksjonistyczne są trudne do obalenia (w rozumieniu podejścia falsyfikacyjnego w nauce). Nie oznacza to jednak, że pewnych, dających się wyprowadzić założeń, nie można poddać weryfikacji. Jak wspomniano wcześniej, Levelt (1983; 1989) wcielił do swojej teorii główną zasadę przerywania mówienia (*MIR*), którą inni badacze podważyli dostarczając przekonujących dowodów na możliwość planowania reperacji jeszcze przed wystąpieniem interupcji (Blackmer i Mitton, 1991; por. Postma i Kolk, 1993). Hartsuiker i Kolk (2001) zaproponowali modyfikację, zgodnie z którą planowanie reperacji oraz interupcja artykulacji stanowią względnie niezależne, równoległe działające procesy. Dzięki opracowanemu modelowi obliczeniowemu, udało im się obronić teorię pętli percepcyjnej przed jeszcze innym zarzutem. Według Postmy (2000) tempo mówienia ma wpływ na czas przechowywania planu fonetycznego w buforze artykulacyjnym, a tym samym na czas w jakim wewnętrzna pętla percepcyjna może wykryć potencjalne błędy i zaszygnalizować konieczność przerwania artykulacji. W przypadku zwiększonego tempa mówienia czas buforowania jest krótszy, przez co błędy powinny być wykrywane z opóźnieniem (jeszcze przez pętlę wewnętrzną lub dopiero przez pętlę zewnętrzną), względem normalnego tempa mówienia, skutkując dłuższymi odcinkami czasu między błędem a interupcją oraz między interupcją a reperacją. Badanie Oomen i Postmy (2000, za: Postma i Oomen, 2005) wykazało jednak efekt odwrotny: mówienie w szybszym tempie skutkowało krótszym czasem między wystąpieniem błędu, a przerwaniem mówienia niż mówienie w normalnym tempie. Według Hartsuikera i Kolka (2001) efekt ten nie jest sprzeczny z założeniami teorii pętli percepcyjnej jeśli uwzględnimy fakt, że mówienie w szybszym tempie

odbija się na całym procesie produkcji mowy (np. na kodowaniu fonologicznym, por. Dell, 1986), a nie tylko na buforowaniu i artykulacji. Symulacje mówienia i monitorowania w szybszym tempie dostarczyły efektów porównywalnych z wynikami badań empirycznych uzyskanych przez Oomen i Postmę (2000, za: Postma i Oomen, 2005), pozwalając pogodzić je z teorią pętli percepcyjnej, bez konieczności włączania do modelu dodatkowego monitora produkcji mowy, który miałby radzić sobie ze wzrostem tempa mówienia (por. Postma, 2000). Co więcej Hartsuiker i Kolk (2001) zauważają, że nie można wykluczyć, iż wraz ze zwiększeniem tempa mówienia również odbiór mowy ulega przyspieszeniu, co pozwoliłoby na zachowanie efektywności pracy wewnętrznej pętli percepcyjnej na niezmiennym poziomie.

Inny problem teorii Levelta (1983; 1989) zauważyli Postma i Kolk (1993), wskazując, że ogranicza ona monitorowanie mowy tylko do kanału słuchowego, gdyż trudno oczekiwać, że system służący do odbioru słyszanej z zewnątrz mowy będzie miał dostęp do dotykowej lub proprioceptywnej (oraz eferentnej) informacji zwrotnej. Choć Levelt (1989; 1999) ani Hartsuiker i Kolk (2001) nie podjęli dalej tej kwestii, najnowsze doniesienia (np. Franken, Liu, Ostry, 2022) pokazują, że w odbiór mowy zaangażowana jest nie tylko modalność słuchowa, ale ważną rolę odgrywają także dane somatosensoryczne (por. Trudeau-Fisette, Ito, Ménard, 2019). Oznacza to, że system odbioru mowy, nie jest „ślepy” na sprzężenie zwrotne z kanałów proprioceptywnego i dotykowego, więc potencjalnie może korzystać z nich również w celu monitorowania produkcji mowy<sup>72</sup>.

Podczas gdy teoria pętli percepcyjnej opiera monitorowanie wyłącznie na systemie odbioru mowy, konkurencyjna teoria koneksjonistyczna MacKaya (1992) zdaje się nie doceniać roli, jaką w monitorowaniu mowy może odgrywać pętla zewnętrzna. Z jednej strony zredukowanie problemu wykrywania błędów do prostego porównywania wskaźników aktywacji węzłów jest atrakcyjne, jednak bez monitorowania zewnętrznego trudno wyjaśnić wyniki badań, w których odsetek wykrytych błędów spada wraz z blokowaniem słuchowej informacji zwrotnej (Postma, 2000). Liczne dane pokazujące, że percepcja własnej mowy ma istotne znaczenie dla monitorowania zewnętrznego (Huettig i Hartsuiker, 2010; Lind i in., 2014) przyczyniły się do tego, że Nozari (2020) kontynuująca koneksjonistyczną tradycję MacKaya i Della, włączyła do swojego modelu założenie o wykrywaniu części błędów przez zewnętrzną pętlę odbioru mowy.

---

72 Dodatkowo badacze (Franken i in., 2022; por. Trudeau-Fisette i in., 2019) sugerują że integrowanie danych z wielu modalności w procesie percepcji mowy jest efektem treningu w aktywnym mówieniu, co jeszcze mocniej akcentuje relacje między systemami produkcji i odbioru mowy, stanowiąc pośrednie wsparcie dla teorii pętli percepcyjnej.

Jak zauważa Roelofs (2020a) teoria monitorowania konfliktu postuluje istnienie dwóch mechanizmów monitorowania: mowa zewnętrzna monitorowana jest przez system odbioru mowy, natomiast monitorowanie wewnętrzne wykrywa błędy poprzez ocenę konfliktu w systemie produkcji mowy. Współcześnie więc dyskusja toczy się nie wokół tego, czy monitorowanie w ogóle opiera się na systemie odbioru mowy, ale czy system ten jest wystarczający, aby efektywnie zapobiegać powstawaniu błędów mowy. Roelofs (2005a; 2020a), kontynuujący tradycję pętli percepcyjnej Levelta, uważa, że monitorowanie tak zewnętrzne, jak i wewnętrzne odbywa się za pośrednictwem systemu odbioru mowy. Jednym z przytaczanych argumentów na rzecz tego twierdzenia jest wrażliwość monitorowania wewnętrznego na efekty obserwowane w badaniach nad odbiorem mowy, np. punkt wyjątkowości wyrazu (*perceptual uniqueness point of words*)<sup>73</sup> (Özdemir, Roelofs i Levelt, 2007). W kontrze, Nozari i in. (2011) zauważają, że monitor pracujący na bazie systemu odbioru mowy powinien być wrażliwy na uszkodzenia tego systemu, tymczasem badania nad pacjentami z zaburzeniami przetwarzania mowy pokazują dysocjację między monitorowaniem a odbiorem mowy. Dokładniej mówiąc, tzw. problem podwójnej dysocjacji wiąże się z obserwacją, że uszkodzony odbiór mowy nie zawsze wiąże się z obniżonymi zdolnościami monitorowania mowy oraz, że pogorszona zdolność wykrywania i poprawiania błędów nie zawsze ma związek z zaburzeniami w odbiorze mowy (Roelofs, 2020a). Hartsuiker i Kolk (2001) twierdzą, że prawidłowy odbiór mowy przy zubożonym monitorowaniu własnej mowy może być efektem uszkodzenia systemu na wyższych piętrach (system odbioru mowy prawidłowo analizuje dostarczone dane językowe, ale zawodzi proces porównywania reprezentacji lub inicjowania reperacji), co nie przeczyłoby teorii pętli percepcyjnej (por. Roelofs, 2020a; 2000b). Dużo bardziej problematyczne są przypadki kliniczne, w których uszkodzony system odbioru mowy nie prowadzi do obniżenia zdolności monitorowania (por. Postma, 2000). Według Nozari i in. (2011) są one dowodem na to, że monitorowanie wewnętrzne nie opiera się (wyłącznie) na systemie odbioru mowy, lecz jest funkcją systemu produkcji mowy oraz kontroli wykonawczej reagującej na wzrost konfliktu w sieci węzłów (Nozari i in., 2011; Nozari, 2020).

---

73 Słyszac pierwsze głoski jakiegoś wyrazu, łatwo pomylić go z innym, gdyż ten sam nagłos współdzielony jest przez wiele wyrazów, np. nagłosowa sylaba [do] występuje w wyrazach domek, dokumenty, doradca, jednak segment [dok] występuje tylko w wyrazie dokumenty. Słyszac sylaby [dokto] wiadomo, że chodzi o wyraz „doktor” (lub pokrewne utworzone od tego samego rdzenia), a nie np. o wyraz doktryna. Punkt wyjątkowości, to miejsce w strukturze wyrazu, po którym jego rozpoznanie jest najłatwiejsze, gdyż od tego momentu nie dzieli on wspólnych fonemów z innymi, niespokrewnionymi wyrazami. Inaczej mówiąc, od tego miejsca dany wyraz różni się od wszystkich innych wyrazów w słowniku umysłowym (Luce, 1986). Efekt unikalności w monitorowaniu mowy został wykazany przez Özdemira i in. (2007). W ich badaniu wewnętrzne monitorowanie fonemów „przyspieszało” po przekroczeniu punktu wyjątkowości wyrazu.

Argumentacja na temat podwójnej dysocjacji wytaczana przeciw wewnętrznej pętli percepcyjnej ma jednak istotne mankamenty. Po pierwsze przykłady pochodzą ze studiów przypadków lub z badań na małych grupach, które trudno uogólnić. Po drugie nie we wszystkich badaniach obserwuje się dysocjację między odbiorem mowy z zewnątrz a monitorowaniem własnego procesu mówienia, przeciwnie niektóre badania pokazują istniejące związki. Np. pacjenci z afazją Wernickego, w której charakterystyczne jest ograniczone rozumienie mowy, są też mniej świadomi popełnianych przez siebie błędów niż pacjenci z afazją Brocka, podczas gdy w afazji przewodzeniowej, pacjenci ze sprawnymi funkcjami odbioru mowy są świadomi swoich błędów i podejmują próby ich reperacji (Roelofs, 2020a). Odmienne wyniki mogą być związane także ze stosowaniem przez badaczy różnych metod do pomiaru zmiennych, np. różne testy do pomiaru rozumienia mowy, w które w różnym stopniu angażują funkcje istotne z punktu widzenia monitorowania<sup>74</sup>. Zresztą sam pomiar monitorowania wewnętrznego stanowi problem metodologiczny, gdyż poza procedurą monitorowania fonemów lub sylab (Morgan i Wheeldon 2003; Wheeldon i Levelt, 1995) w badaniach nad wykrywaniem i poprawianiem błędów stosuje się miary pośrednie związane z ujawnianiem ukrytych reperacji. Jak już wspomniano ukryte reperacje są niejednoznaczne i trudne do klasyfikacji (Postma, 2000), gdyż w wypowiedzi najczęściej nie uzewnętrznia się ewidentny błąd, a obserwowane zakłócenia płynności mowy mogą mieć inne przyczyny niż wykrycie błędu w planie przedartykulacyjnym (por. Engelhardt i in., 2013). Z tego względu nie ma zgody, co do niezawodnych wskaźników ukrytych reperacji i działania monitorowania wewnętrznego, np. jedni badacze jako przykłady ukrytych reperacji klasyfikują powtórzenia fraz, słów lub ich części (Oomen i in., 2005), podczas gdy inni (Nozari i in., 2011)<sup>75</sup> decydują się nie klasyfikować powtórzeń jako przejawów wewnętrznego wykrywania błędów. Podobne wątpliwości dotyczą statusu pauz wypełnionych (Hartsuiker, Kolk i Martensen, 2005).

Ostatnim, być może najważniejszym zastrzeżeniem do argumentu podwójnej dysocjacji jest jego ograniczone odniesienie do właściwego rozumienia teorii pętli percepcyjnej. Według Levelta (1989; 1992) cele słuchania wypowiedzi innych osób oraz monitorowania swojej własnej mowy są odmienne. Słuchacz musi przede wszystkim wydobyć ze strumienia mowy sens słyszanego tekstu, podczas gdy nadawca ocenia poprawność generowanego komunikatu i jego zgodność z intencją komunikacyjną. Korzystanie przez procesy rozumienia i monitorowania z tego samego systemu nie oznacza

74 Często przywoływane w kontekście podwójnej dysocjacji badanie Nickelsa i Howarda (1995; za: Roelofs, 2020a) doczekało się dwóch zupełnie odmiennych interpretacji (por. Roelofs, 2005a; Nozari i in., 2011), w zależności od tego, które miary do oceny odbioru mowy zostały uwzględnione w analizie.

75 Jednocześnie ci sami badacze (Nozari i in., 2011) sprawność monitorowania wewnętrznego mierzą na podstawie wszystkich wykrytych błędów (niezależnie od poprawności reperacji), co rodzi zarzut o brak wyraźnego rozgraniczenia między monitorowaniem wewnętrznym a zewnętrznym (Roelofs, 2020a).

korzystania z niego przez oba procesy w taki sam sposób. W podobnej linii przebiega argumentacja Roelofsa (2020a; 2000b), broniąca koncepcji automonitorowania przez odbiór mowy. Według niego skuteczne słuchanie i rozumienie innych osób zależy od tego, jak dobrze pacjent (ale też zdrowy słuchacz) jest w stanie przetwarzać zewnętrzny sygnał, począwszy od cech akustycznych przez fonemy, jednostki leksykalne do pojęć. Natomiast podczas mówienia selekcjonowane pojęcia, elementy leksykalne i fonemy aktywują odpowiednie reprezentacje w systemie odbioru mowy, przez co monitorowanie przebiegu mówienia może być mniej wrażliwe na uszkodzenia, które z kolei zaburzają procesy rozumienia informacji językowej. Z tego względu teoria pętli percepcyjnej dopuszcza obserwowanie stosunkowo skutecznego monitorowania własnej mowy przy jednoczesnym obniżeniu rozumienia wypowiedzi innych osób, a argument podwójnej dysocjacji nie falsyfikuje założenia o wewnętrznej pętli percepcyjnej.

Wracając do porównania modeli monitorowania mowy, teoria pętli percepcyjnej pozostaje najbardziej wpływową propozycją wśród poznawczych podejść do kontroli procesu mówienia. Zakłada ona, że osoba mówiąca monitoruje swoje wypowiedzi poprzez pętlę zewnętrzną (na drodze słuchowego odbioru własnej mowy) oraz pętlę wewnętrzną mającą dostęp do reprezentacji wyrazów fonologicznych. Obie pętle umożliwiają porównanie efektów programowania mowy z przedwerbalnym przekazem za pośrednictwem systemu odbioru mowy. Konkuruje z nią teoria monitorowania konfliktu Nozari i in. (2011; Nozari, 2020) także przyjmuje, że monitorowanie zewnętrzne (monitorowanie mowy artykułowanej) dokonuje się podobnie, jak rozumienie komunikatów językowych innych ludzi. Zasadnicza różnica dotyczy więc mało uchwytne mechanizmu monitorowania wewnętrznego. Zdaniem Nozari (2018) monitor, który miałby dostęp jedynie do mowy wewnętrznej, stanowiącej końcowy produkt przedartykulacyjnego programowania mowy, nie byłby w stanie efektywnie poradzić sobie z szybkim wykrywaniem błędów na wczesnych etapach produkcji mowy. Według niej systemem, który w najlepszy sposób może poradzić sobie z problemami powstającymi w procesie mówienia jest system produkcji mowy. Wniosek ten wynika z przyjętych założeń na temat działania sieci leksykalnej w procesie generowania mowy, w której występują dwukierunkowe oddziaływania między poziomami węzłów. Dzięki informacjom zwrotnym między węzłami, możliwe jest więc rejestrowanie nietypowych wzorców aktywacji, które mogą zaalarmować nadawcę o potencjalnym błędzie. Teoria Levelta i in. (1999), która zakłada jednokierunkowe rozprzestrzenianie się aktywacji, nie dopuszcza mechanizmu zwrotnej aktywacji węzłów wyższego rzędu, przez węzły z poziomu niższego, stąd informacja zwrotna o postępie w produkcji mowy dostępna jest dopiero na

końcu tego procesu<sup>76</sup>, a jej weryfikacja wymaga zaangażowania systemu odbioru mowy. W rozwinięciu modelu zaproponowanym przez Roelofsa (2005a; 2020a) wykrywanie błędów wymaga operacji poznawczych porównujących reprezentacje wybrane w procesie mówienia z reprezentacjami wybranymi (aktywowanymi) w procesie rozumienia<sup>77</sup> i oceniających rozbieżności między nimi (por. Roelofs, 2020b). Na przykład błędy selekcji leksykalnej mogą zostać wykryte przez sprawdzenie tego, czy lemma rozpoznanego słowa odpowiada pojęciu leksykalnemu aktywowanemu na etapie rozpoczynania produkcji. Za sprawdzenie tej zgodności odpowiada mechanizm weryfikacji zakładany przez model WEAVER++ (Levelt i in., 1999).

Kluczowe różnice w obu teoriach wynikają więc z odmiennych założeń dotyczących organizacji i przebiegu produkcji mowy. Lind i Hartsuiker (2020, s. 2) zauważają, że problem w uzyskaniu zgodności między tymi propozycjami teoretycznymi odsyła do bardziej podstawowego problemu: „jak najlepiej modelować system produkcji języka”. W związku z tym, że rozstrzygnięcie między modelami produkcji mowy nie jest celem niniejszej rozprawy (charakterystyka najważniejszych modeli została przedstawiona w rozdziale 2), w dalszej części tekstu nie będą podkreślane rozbieżności między różnymi teoriami monitorowania procesu mówienia, a uwypuklone zostaną pewne punkty wspólne.

Zarówno teoria monitorowania konfliktu Nozari i in. (2011) oraz teoria pętli percepcyjnej (w wersji zaproponowanej przez Roelofsa, 2005a; 2020a) wiążą kontrolę procesu mówienia z ogólnymi mechanizmami kontroli wykonawczej. Jak wspomniano wcześniej, Nozari uważa, że konflikt jest wykrywany przez ogólny system kontrolny (niespecyficzny dla funkcjonowania językowego) zlokalizowany w przednim zakręcie kory obręczy (ACC). Struktura ta jest wiązana z kontrolą wykonawczą czynności poprzez monitorowanie konfliktów powstających w wyniku rozbieżności między zakładanymi a osiąganymi celami, np. w wyniku popełnienia błędu (Fąfrowicz i Marek, 2008). Według Nozari (2018) te ogólne mechanizmy (*domain-general mechanisms*) mogą działać na reprezentacjach specyficznych dla danej domeny (*domain-specific representations*), np. języka, tworząc wyspecjalizowane pętle monitorowania i kontroli służące wykrywaniu

---

76 Warto odnotować tu pewną nieścisłość. Pierwotnie Levelt (1989) zakładał, że pętla wewnętrzna operuje na planie fonetycznym wypowiedzi, stanowiącym końcowy produkt programowania mowy tuż przed rozpoczęciem artykulacji. Po aktualizacji teorii i dodaniu twierdzenia o dostępie w monitorowaniu do reprezentacji fonologicznej rozszerzono tym samym zakres monitorowania na „produkt pośredni” (por. Schiller, 2005; Hartsuiker i in., 2007). Ta nieścisłość byłaby problemem dla teorii modularnej, zakładającej hermetyczność i nieprzenikliwość poznawczą modułów. W wersji teorii z 1999 roku (Levelt i in., 1999; Levelt, 1999) kodowanie morfofonologiczne ma charakter sieciowy, nie musi więc spełniać kryteriów modularności (por. Hartsuiker i Moors, 2017)

77 *Org. errors are detected by comparing representations selected for production with representations selected in comprehension* (Roelofs, 2020a, s. 3).

błędów w mówieniu. Co istotne, również podejście reprezentowane przez Roelofsa (2020a) akcentuje powiązanie monitorowania procesu mówienia z niespecyficznymi mechanizmami kontroli czynności. W przeciwieństwie jednak do propozycji Nozari i in. (2011), Roelofs uważa, że główną rolą przedniego zakrętu kory obręczy (ACC) i „innych obszarów czołowych” jest sprawowanie i regulowanie kontroli wykonawczej, a nie monitorowanie konfliktu<sup>78</sup>.

W kolejnych podrozdziałach opisane zostaną tzw. funkcje wykonawcze (*executive functions*), które tradycyjnie wiązane są ze wskazywanymi tutaj strukturami anatomicznymi mózgu (obszary czołowe wraz z ACC) oraz ich potencjalne związki z kontrolą procesu mówienia.

## 3.2 Funkcje wykonawcze

Termin „funkcje wykonawcze” lub „zarządcze” (*executive functions*<sup>79</sup>) jest używany do opisu procesów odpowiadających za samoregulację i sterowanie celowym działaniem. Diamond (2013) definiuje je jako grupę odgórnych procesów umysłowych (*top-down mental processes*) potrzebnych w działaniu wymagającym skupienia uwagi, gdy poleganie na instynkcie lub wyuczonych, zautomatyzowanych nawykach byłoby niewłaściwe, niewystarczające lub niemożliwe. Jak zauważa Jodzio (2008, s. 32) funkcje wykonawcze „w sposób »poznawczy« przygotowują podmiot do (...) działania oraz pozwalają ocenić jego skuteczność”. Innymi słowy, funkcje wykonawcze integrują i organizują inne procesy poznawcze, co przekłada się na poziom zorganizowania i kontroli celowych działań (Józefacka-Szram, 2014). O ich znaczeniu dla prawidłowego zarządzania zachowaniem świadczą liczne badania, wiążące poziom przejawianej kontroli wykonawczej z takimi zmiennymi jak zdrowie fizyczne i psychiczne, jakość życia, osiągnięcia edukacyjne i zawodowe (Diamond, 2013).

Konstrukt ten obejmuje złożoną grupę procesów kontrolnych, wśród których najogólniejszy podział pozwala rozróżnić funkcje „zimne” i „gorące”. Te pierwsze odgrywają istotną rolę, wówczas gdy osoba „rozwiązuje względnie abstrakcyjne, oderwane od kontekstu problemy, w których uzyskanie określonego wyniku nie wiąże się z istotnymi

78 Różnice w ocenie udziału kory przedniego zakrętu obręczy (ACC) wynikają z odmiennej interpretacji funkcji tej struktury w wykrywaniu błędów. Podczas gdy Nozari (2020) uważa, że ACC monitoruje poziom konfliktu, Roelofs (2020) jest zdania, że ACC reaguje na sygnał o błędzie wykrytym przez inne obszary kory. Fąfrowicz i Marek (2008, s. 157) uważają, że wskazanie funkcji poznawczych specyficznych dla przedniego zakrętu obręczy stanowi „jedno z większych wyzwań neurokogniistyki”. Na razie jednak zgoda badaczy kończy się na dość ogólnie zarysowanym przyjęciu, że ACC jest systemem egzekucyjno-kontrolnym.

79 W literaturze przedmiotu można spotkać także zamiennie określenia „kontrola wykonawcza” (*executive control*) lub „kontrola poznawcza” (*cognitive control*) (Diamond, 2013).

konsekwencjami dla jednostki” (Putko, 2011, s. 74). Zalicza się do nich takie umiejętności jak planowanie, hamowanie, elastyczność, pamięć roboczą, czy monitorowanie wykonania zadania (Peterson i Welsh, 2014). Funkcje gorące uruchamiane są natomiast w sytuacji postawienia jednostki przed wyborem szybkiej gratyfikacji lub długoterminowej nagrody, czyli „są zaangażowane w rozwiązywanie problemów związanych z regulacją afektu i motywacji” (Putko, 2011, s. 74).

Z punktu widzenia celów niniejszej pracy istotnym jest, aby zawęzić grupę omawianych procesów. Jakkolwiek funkcje „gorące” mogą mieć udział w procesie mówienia, wpływając szczególnie na jego stronę motywacyjną, związków monitorowania i kontroli produkcji mowy z procesami kontroli poznawczej należałoby szukać raczej w obrębie funkcji „zimnych” (por. Shao i in., 2013). Tradycyjnie też funkcje „gorące” utożsamiane są bardziej z pojęciem samoregulacji (*self-regulation*), a funkcje „zimne” z kontrolą wykonawczą w przyjętym tu rozumieniu (Diamond, 2013; por. Singer i Bashir, 1999). Dlatego w dalszej części opracowania właśnie tej grupie procesów poświęcona zostanie uwaga. Warto jeszcze zaznaczyć, że obie grupy funkcji podlegają rozwojowi w ciągu dzieciństwa i adolescencji, osiągając najwyższy poziom sprawności w okresie dorosłości. Jednocześnie funkcje „gorące” rozwijają się wolniej i wzrost ich sprawności obserwuje się później niż funkcji „zimnych” (Prencipe, Kesek, Cohen, Lamm, Lewis i Zelazo, 2011).

Jak już wspomniano pojęcie funkcji wykonawczych (w dalszej części termin ten będzie używany na oznaczenie jedynie funkcji „zimnych”, por. Diamond, 2013) jest konstruktem złożonym. Choć niektórzy autorzy postulują jedność, czy jednorodność funkcji wykonawczych, zdecydowana większość badaczy podkreśla ich wieloaspektową naturę (Huizinga, Dolan, i van der Molen, 2006). Według Jodzio (2008) mówić można o wieloczynnikowej organizacji i heterogenicznym charakterze procesów, które zbiorczo określane są tym terminem. Wiąże się to z wielością i różnorodnością przejawów kontroli wykonawczej w działaniu człowieka (Jankowski, 2012). Funkcje wykonawcze leżą u podstaw opóźniania lub powstrzymywania się od impulsywnych reakcji, hamowania czynności już rozpoczętych, selektywnego kierowania uwagą, ustalania celów, planowania i inicjowania czynności, rozwiązywania problemów, trzymania się reguł i elastycznego przełączania między nimi (Singer i Bashir, 1999; Józefacka-Szram, 2014). Neuronalnym podłożem funkcji wykonawczych są struktury należące do płatów czołowych, ze szczególnym uwzględnieniem grzbietowo-bocznej kory przedczołowej (Jodzio, 2008; Rajtar i in., 2014). Osoby, u których struktury te ulegają uszkodzeniu przejawiają trudności w planowaniu, hamowaniu reakcji, zapamiętywaniu i elastycznym przechodzeniu między celami (Niendam i

in., 2012). Jednocześnie podobne zaburzenia w planowaniu, inicjowaniu oraz kontroli podjętych działań obserwuje się także u pacjentów z uszkodzeniami struktur podkorowych (Jodzio i in., 2012). Oznacza to, że obwody neuronalne powiązane z funkcjami wykonawczymi nie ograniczają się tylko do kory przedczołowej, ale mogą obejmować także struktury podkorowe (Jodzio, 2008)<sup>80</sup>. Na przykład, w zależności od przyjętej konceptualizacji funkcji hamowania i wynikającej z tego metodologii pomiaru w badaniach neuroobrazowania, oprócz struktur kory przedczołowej wykazuje się udział kory zakrętu obręczy, jąder podstawnych, czy kory przedruchowej (Nigg, 2000; Xue i in., 2008). Metaanaliza prawie 200 badań przeprowadzonych łącznie na ponad 2,5 tys. osób w wieku dorosłym wykazała, że w wykonanie większości zadań tradycyjnie wiązanych z funkcjami wykonawczymi angażuje nadrzędną sieć kontroli poznawczej, która obejmuje obszary czołowe, ciemieniowe, zakręt obręczy, i struktury podkorowe (Niendam i in., 2012; por. Pąchalska, Kaczmarek i Kropotov, 2014).

Tak złożona sieć może obsługiwać wiele subprocesów i manifestować się na różne sposoby w ludzkim działaniu, dlatego też powstało wiele różnych klasyfikacji funkcji wykonawczych, odmiennie wyodrębniających prostsze komponenty z tego złożonego konstruktu (Jodzio, 2008; Jankowski, 2012). Według Diamond (2013) panuje jednak względny konsensus na temat zasadności wyróżnienia trzech podstawowych funkcji wykonawczych: hamowania (*inhibition*), elastyczności poznawczej (*cognitive flexibility*) i pamięci roboczej (*working memory*)<sup>81</sup>. Źródłem wyodrębnienia tych własne funkcji są analizy czynnikowe przeprowadzone na wynikach uzyskiwanych przez osoby badane w różnych zadaniach mierzących funkcje wykonawcze (Miyake i in. 2000; Lehto i in., 2003). Analizy te wykazują istnienie trzech względnie niezależnych (choć korelujących ze sobą) komponentów funkcji wykonawczych, które w różnym stopniu odpowiadają za poziom wykonania poszczególnych testów mierzących ich poziom. W dalszej części rozdziału, każda z tych trzech funkcji wykonawczych zostanie dokładniej scharakteryzowana, zwłaszcza że w

---

80 Jodzio (2008) opisując mózgową organizację funkcji wykonawczych, zwraca uwagę na liczne powiązania kory przedczołowej z prążkowiec, którego funkcja – wbrew temu co dawniej sądzono – nie ogranicza się tylko do kwestii wyłącznie ruchowych (jako części układu pozapiramidowego), ale może mieć znaczenie dla mowy, uwagi czy pamięci proceduralnej. Szczególnie interesujący (w kontekście kontroli procesu mówienia) jest domniemany udział prążkowiec w sekwencyjnej organizacji czynności, która to z kolei może być ważna np. w przetwarzaniu gramatycznych konstrukcji zdań. Warto w tym miejscu zaznaczyć, że część połączeń między korą przedczołową (część przyśrodkowa), a prążkowiec (jądro półleżące) odbywa się za pośrednictwem kory przedniego zakrętu obręczy (ACC).

81 Poza podziałem na podstawowe funkcje wykonawcze Diamond (2013) proponuje wyróżnienie także trzech funkcji wyższego stopnia (rozumowanie, rozwiązywanie problemów, planowanie), które nadbudowywane są na tych pierwszych.

obrębie poszczególnych funkcji daje się na bazie analizy zmiennych latentnych wyodrębnić kolejne składowe (Huizinga i in., 2006; McAuley i White, 2011)<sup>82</sup>.

### **3.2.1 Kontrola hamowania**

Istotną cechą pozwalającą zaklasyfikować dany proces do grupy funkcji wykonawczych jest jego odpowiedzialność za świadomą kontrolę myśli i działań (Putko, 2011). Jednocześnie wiele procesów poznawczych oraz czynności zewnętrznych przebiega poza pełną świadomością. Jest to podyktowane ekonomią poznawczą, gdyż procesy zautomatyzowane są mniej podatne na zakłócenia, przez co zużywają także mniej zasobów poznawczych. Reagowanie w nawykowy, automatyczny sposób nie zawsze jest jednak korzystne, gdyż okoliczności mogą wymagać innego sposobu zachowania się lub odroczenia reakcji. Poza tym procesy zautomatyzowane mogą stać się źródłem interferencji dla czynności nieautomatycznej (Nęcka i in. 2006). Umiejętność zahamowania reakcji, czyli powstrzymanie się przed działaniem impulsywnym lub jego odroczenie, nazywane jest hamowaniem lub kontrolą hamowania (Brzezińska i Nowotnik, 2012; Diamond, 2013). Na to pojęcie składają się trzy powiązane ze sobą procesy: zahamowanie dominującej reakcji na bodźce, przerwanie trwającej reakcji w celu jej modyfikacji, a także ochrona działania celowego przed zakłócającym wpływem dystraktorów (Woźniak-Prus i Gambin, 2011). Pierwsze dwa procesy określane są mianem hamowania behawioralnego (*behavioral inhibition*), które odpowiada za powstrzymanie lub przerwanie (w trakcie trwania) automatycznej, nawykowej lub wcześniej przygotowanej reakcji (Nigg, 2000). Ochrona działań celowych przed wpływem dystraktorów poprzez zarządzanie uwagą selektywną (*selective attention*) nazywana jest z kolei kontrolą interferencji (*interference control*) (Diamond, 2013). Ten typ hamowania wiąże się z dowolnym kierowaniem uwagą, poprzez blokowanie odbioru pewnych bodźców, a koncentrowanie się na tych, które związane są z podjętymi celami i intencją. Podobnie bodźce wewnętrzne (myśli, wspomnienia) mogą stanowić źródło interferencji dla działania celowego, stąd aktywne hamowanie poznawcze (*cognitive inhibition*) także zaliczane jest do przejawów hamowania wykonawczego (Nigg, 2000). Z racji tego, że hamowanie poznawcze ma

---

82 Problem z jasnym zdefiniowaniem i kompletnym wyliczeniem poszczególnych funkcji wykonawczych przypomina kłopoty z określeniem, czym jest inteligencja (por. Nęcka, 2003). Obserwowany jest bowiem prymat badań empirycznych nad refleksją teoretyczną w tej dziedzinie, przez co funkcje wykonawcze utożsamiane są z tym, co badają określone testy, a do „stosowanej już metody próbuje się dobudować fundament teoretyczny” (Jodzio, 2008, s. 64-66). Tym samym trudno jest precyzyjnie opisać dany proces bez odsyłania do zadania badającego jakiś jego aspekt. W związku z tym przedstawiona w dalszej części pracy charakterystyka funkcji wykonawczych połączona będzie ze wskazaniem typowego zadania służącego do jej pomiaru lub problemu w jakim się objawia.

szczególnie ochronne znaczenie dla pamięci roboczej, jego pomiar bardziej koresponduje z miarami pamięci roboczej niż innymi przejawami hamowania (Diamond, 2013).

Shao i in. (2013). wskazują, że powyższe rodzaje hamowania można sprowadzić do dwóch kategorii: hamowanie selektywne i nieselektywne. Hamowanie selektywne wiąże się z reagowaniem w warunkach konkurencji między możliwymi reakcjami, wynikającej z zewnętrznych dystraktorów. Hamowanie nieselektywne wiąże się natomiast z powstrzymaniem się od jakiegokolwiek niepożądanego reakcji, bez konieczności wyboru spośród konkurencyjnych (pożądane jest nie reagować).

Według Shao i in. (2013) hamowanie nieselektywne ujawnia się w zadaniach:

1. Zadanie Go/no-go (*Go/no-go task*) – zadanie polega na wykonywaniu odpowiedniej reakcji (np. prosta reakcja motoryczna poprzez naciśnięcie klawisza) w odpowiedzi na dany bodziec i powstrzymaniu tej reakcji w momencie ujżenia innego bodźca.
2. Zadanie sygnału stop (*Stop signal task*) – w tym zadaniu osoba badana widzi dwa rodzaje bodźców, na które reaguje (reakcja wyboru) poprzez naciskanie odpowiedniego przycisku w momencie wyświetlenia danego bodźca. W części prób prezentowany jest również bodziec słuchowy sygnalizujący konieczność zahamowania reakcji. Czas emitowania sygnału dźwiękowego dopasowuje się do indywidualnego tempa hamowania osoby badanej.

W obu rodzajach zadań wymagane jest powstrzymanie się od reakcji w określonych okolicznościach, bez konieczności wybierania właściwej reakcji. Oba typy zadań Nigg (2000) przypisuje do hamowania behawioralnego. Z kolei hamowanie selektywne jest obecne w zadaniach związanych z kontrolą interferencji (Shao i in, 2013; por. Nigg, 2000):

1. Test Stroopa (*Stroop Test*) – osobom badanym wyświetlane są na ekranie kolorowe napisy, odnoszące się do nazw kolorów. Zadaniem osób badanych jest nazywanie kolorów czcionki (lub ich określanie w inny, niewerbalny sposób – np. wybór przypisanego danemu kolorowi klawisza na klawiaturze), a ignorowanie treści napisów. W warunku zgodności (kiedy nazwy kolorów pokrywają się z kolorem czcionki) czas reakcji jest krótszy niż w warunku niezgodności (nazwa koloru zapisana jest czcionką o innym kolorze), co jest miarą siły interferencji.
2. Zadanie z flankerami (*Flanker task*) – zadaniem badanych jest identyfikowanie kierunku figury centralnej prezentowanej w dwóch rodzajach orientacji względem figur sąsiadujących: orientacja zgodna (<<<<<<) i niezgodna (<<><<). Podobnie jak w teście Stroopa, średni czas reakcji oraz liczba popełnianych błędów różnicują próby w obu warunkach. Oprócz hamowania reakcji zadanie wymaga uwagi selektywnej i jest

czułe na zmiany rozwojowe w późnym dzieciństwie i adolescencji (Józefacka-Szram, 2014).

### **3.2.2 Pamięć robocza**

Pamięć robocza (*working memory*) to proces umysłowy odpowiedzialny za ograniczone w czasie przechowywanie informacji związanej z wykonywanym zadaniem. Informacja ta może być nie tylko przechowywana, ale również przetwarzana w celu rozwiązania zadania, co pociąga za sobą zwiększenie obciążenia zasobów pamięciowych (Piotrowski i in., 2009). Funkcję pamięci roboczej odnaleźć można w wykonywaniu obliczeń matematycznych „w pamięci”, w wyobrażeniowym porządkowaniu różnych elementów (np. tworząc listę zadań lub zakupów) w rozważaniu różnych alternatyw, przekładaniu instrukcji na plan działania oraz aktualizowaniu tego planu przez włączanie nowych informacji itp. Tym co łączy wspomniane wyżej czynności umysłowe, to przetwarzanie informacji, które przestały być dostępne zmysłowo, łącząc w sensownym działaniu zdarzenia minione z tym, co dopiero nastąpi (Diamond, 2013).

Konstrukt ten ma swoje źródła w pojęciu pamięci krótkotrwałej (*short-term memory*), jednak nie jest jej synonimem. Jest używany przede wszystkim, aby zaznaczyć, że zapamiętywana chwilowo informacja nie jest tylko biernie przechowywana w tymczasowych magazynach, ale również aktywnie przetwarzana dla celów związanych z bieżącym zadaniem. Z tego względu badacze rozdzielają procesy przechowywania informacji w pamięci roboczej (*working memory storage*) od procesów odświeżania lub aktualizowania (*updating*), co znajduje swoje uzasadnienie w analizach czynnikowych (McAuley i White, 2011). Miyake i in. (2000) uważają odświeżanie za kluczowy element pamięci roboczej, stanowiący – obok hamowania i przełączania (giętkości poznawczej) – fundament funkcji wykonawczych. Warto jednak odnotować, że większość modeli pamięci roboczej nie redukuje jej tylko do tej jednej funkcji. Na przykład jedna z najbardziej znanych teorii pamięci roboczej, zaproponowana przez Baddeleya (1998) zakłada istnienie trzech<sup>83</sup> składników pamięci roboczej:

1. Pętla fonologiczna odpowiada za utrzymywanie w pamięci śladów pamięciowych o charakterze werbalnym dzięki mechanizmowi powtarzania (inna nazwa: pętla artykulacyjna).

---

83 Nowszą propozycją Baddeleya zawiera dodatkowy (czwarty składnik), tzw. bufor epizodyczny (Orzechowski, 2012).

2. Szkicownik wzrokowo-przestrzenny pełni rolę tymczasowego magazynu (bufora) dla reprezentacji wizualnych.
3. Centralny wykonawca – to mechanizm kontrolny odpowiedzialny za (1) bieżące przetwarzanie informacji zgromadzonych w magazynach (buforach), (2) nadzór nad przetwarzaniem informacji poprzez ich selekcję (funkcja zbliżona do kontroli hamowania) oraz (3) koordynację buforów i rozdzielanych między nie zasobów poznawczych, co wymaga aktywnego aktualizowania ich zawartości (Orzechowski, 2012).

Propozycja ta łączy w jeden system funkcjonalny dwa na ogół odróżniane od siebie rodzaje pamięci roboczej tj. pamięć werbalną (pętla fonologiczna) oraz pamięć przestrzenną (szkicownik wzrokowo-przestrzenny). Ze względu jednak na postulowany, wielozadaniowy mechanizm kontrolny (centralny wykonawca), któremu Baddeley przypisuje tak wiele zadań wykonawczych, teoria ta znacznie rozszerza pojęcie pamięci roboczej, czyniąc z niej w zasadzie synonim funkcji wykonawczych (Diamond, 2013; por. Jodzio, 2008). W dalszej części rozprawy termin „pamięć robocza” będzie odnoszony jedynie do funkcji przechowywania informacji i operowania nimi zgodnie z bieżącymi celami, bez wchodzenia w zakres znaczeniowy innych funkcji wykonawczych.

Podstawowe metody badawcze i diagnostyczne pamięci roboczej wiążą się z zadaniami, w których konieczne jest zapamiętywanie cyfr, słów lub ułożenia przedmiotów w przestrzeni. Zadanie w którym należy dane elementy zapamiętać i odtworzyć mierzy zakres pamięci i występuje w wersji odtwarzania wprost i wstecz. Odtwarzanie wprost angażuje przede wszystkim procesy przechowywania informacji w pamięci i tradycyjnie były stosowane do pomiaru pojemności pamięci krótkotrwałej. Odtwarzanie wstecz z kolei wymaga uszeregowania elementów w odwrotnej kolejności niż zostały one podane, angażuje więc aktywne przetwarzanie zapamiętanej informacji. Przykładami zadań stosowanych w powyżej wskazany sposób są:

1. Powtarzanie cyfr skali inteligencji Wechslera (wersja dla dorosłych i dla dzieci) – osobom badanym prezentuje się zestawy cyfr o rosnącej liczbie elementów. W warunku odtwarzania wprost, po usłyszeniu zestawu cyfr, osoba badana powtarza je w kolejności, w jakiej zostały one zaprezentowane, natomiast w warunku odtwarzania wstecz, należy odtworzyć listę w odwrotnej kolejności.
2. Klocki Corsiego – osobom badanym prezentuje się zestaw sześciennych klocków, wskazując na nie w pewnym porządku. Zadaniem uczestnika badania jest odtworzenie tej (lub odwrotnej) sekwencji wskazań poszczególnych klocków.

3. N-wstecz (*n-back*) – procedura polega na tym, że osobie badanej przedstawia się kolejno bodźce (np. litery, figury geometryczne). Jej zadaniem jest stwierdzić, czy aktualnie prezentowany bodziec jest taki sam, czy inny niż bodziec występujący tuż przed nim (1-wstecz) albo poprzedzający go o 2 lub 3 (2-wstecz lub 3-wstecz) pozycje w sekwencji. W zadaniu tym konieczne jest więc ciągle uaktualnianie informacji przetwarzanych przez pamięć roboczą.

Spośród wskazanych powyżej zadań „n-wstecz” traktowane jest jako właściwe do pomiaru odświeżania pamięci roboczej. Bardziej zaawansowane metody, angażujące dodatkowe funkcje wykonawcze wiążą się z obciążaniem zasobów poznawczych tak, aby wystąpiła konkurencja o nie między poszczególnymi komponentami pamięci (Piotrowski, Stettner, Orzechowski, i Balas, 2009).

### **3.2.3 Giętkość poznawcza**

Giętkość poznawcza definiowana jest jako umiejętność elastycznej adaptacji do wymagań otoczenia poprzez sprawne przełączanie się między kryteriami poprawności wykonywanych zadań (Brzezińska i Nowotnik, 2012; Jankowski, 2012). Na oznaczenie tej funkcji używa się także terminów elastyczność poznawcza oraz zmiana nastawienia lub przełączanie (*shifting*) (Gambin i Woźniak-Prus, 2011). Wszystkie te określenia uwypuklają różne aspekty wykorzystania giętkości poznawczej w działaniu (dostosowanie do zmiennych okoliczności, modyfikacja planów, płynne przechodzenie między regułami wyznaczającymi poprawność działania), ale łączy je podkreślanie zdolności do wyjścia poza sztywny, stereotypowy sposób reagowania. Elastyczność poznawcza umożliwia skuteczne dostosowanie się do wymagań lub priorytetów, które uległy zmianie, ale też pozwala przyznać się do błędu lub skorzystać z niespodziewanych, nagłych okazji (Diamond, 2013). Dzięki giętkości poznawczej skutecznie uczymy się wyjątków gramatycznych, rozwiązujemy problemy niejednoznaczności oraz rozumiemy metaforyczny sens wypowiedzi (Filipiak, 2018).

Chociaż przełączanie wyróżnia się w pomiarze funkcji wykonawczych, jako odrębny komponent (Miyake i in. 2000; Lehto i in., 2003), to jednak do pewnego stopnia jego funkcjonalność współzależy od efektywności hamowania i pamięci roboczej. Według Diamond (2013), aby zmienić perspektywę (przestrzenną lub interpersonalną) potrzeba zahamować wcześniejszy sposób oceny sytuacji i aktywować w pamięci roboczej nową perspektywę. To samo dotyczy rozwiązywania innych problemów, w których zmiana sposobu

działania, wybór innej metody i wyjście poza schemat postępowania są fundamentem powodzenia. Z tego względu trudno jest znaleźć zadanie badawcze lub diagnostyczne, które pozwalałoby zmierzyć poziom elastyczności bez angażowania innych funkcji wykonawczych<sup>84</sup>

Do badania giętkości poznawczej używa się metod polegających na przełączaniu się między regułami lub przechodzeniu między zadaniami. Jednym z najstarszych narzędzi angażujących elastyczność poznawczą jest Test Sortowania Kart z Wisconsin (*Wisconsin Card Sorting Task, WCST*), będący klasycznym testem funkcji kory przedczołowej (Diamond, 2013). Zadaniem osoby badanej jest dopasowywanie zmieniających się kart do jednej z 4 prezentowanych (niezmiennie) kart wzorcowych. Karty różnią się zamieszczonymi na nich kształtami figur (gwiazdki, koła, trójkąty i krzyżyki), liczbą figur (od 1 do 4) i ich kolorem (niebieski, żółty, czerwony, zielony). Reguła sortowania nie jest podana jawnie, a uczestnik musi ją określić samodzielnie, na podstawie otrzymywanej po każdej reakcji informacji zwrotnej o poprawności dokonanego wyboru. Zmiana reguły następuje bez wiedzy osoby badanej. Zmiennymi zaleźnymi są tu liczba prawidłowo (zgodnie z regułą) dopasowanych kart oraz liczba błędów perseweracyjnych (sortowanie zgodnie z kryterium, które przestało obowiązywać). Jak wspomniano, funkcja giętkości poznawczej ujawnia się w zadaniach angażujących inne funkcje, dlatego Test Sortowania Kart z Wisconsin bywa stosowany także jako miara pamięci roboczej (Piotrowski, Stettner, Wierzchoń, Balas i Bielecki, 2009) lub poziomu sprawności funkcji wykonawczych w ogóle.

Do innych miar służących badaniu elastyczności służą testy przełączania zadań (*task-switching paradigms*), które wymagają rozwiązywania dwóch podobnych zadań, różniących się jakąś zasadą (Diamond, 2013). Osoba badana jest instruowana, aby naprzemiennie rozwiązywać kolejne etapy obu zadań poprzez płynne przechodzenie od pracy nad jednym zadaniem, do postępu w drugim. Wymaga się więc od niej ciągłej zmiany kryterium wedle, którego wybierana jest prawidłowa odpowiedź, przy czym wybór ten jest ograniczony do dwóch konkurencyjnych wobec siebie zasad. Przykładem takiego zadania jest Test Łączenia Punktów Reitana A i B (Trail Making Test A&B) (Piotrowski in., 2009). W zadaniu TMT A, osoba badana ma w jak najszybszym czasie połączyć linią ciągłą rozmieszczone na kartce lub ekranie punkty ponumerowane od 1 do 25. W zadaniu TMT B chodzi natomiast o to, żeby połączyć punkty rosnąco według porządku alfabetycznego oraz według wzrastającej wartości liczbowej, stosując oba kryteria naprzemiennie (raz wybierany jest punkt z liczbą, a raz z

---

84 Problem zanieczyszczenia, kontaminacji wyników testów badających funkcje wykonawcze jest jednym z powodów podejmowania analiz czynnikowych wyników uzyskiwanych w zdaniach mierzących w założeniu te same funkcje wykonawcze (Miyake i in., 2000).

literą: 1-A-2-B-3-C... itd.) Podobnie jak WCST zadanie to wymaga elastyczności, ale też angażuje funkcje hamowania (stłumienie tendencji do wybierania punktów według powtarzalnego kryterium) oraz pamięć roboczą (utrzymywanie w pamięci zasady postępowania nie jest bardzo wymagające przy jedynie dwóch kryteriach wyboru punktów, ale konieczne jest ciągle aktualizowanie informacji o tym, które z kryteriów jest w danym momencie obowiązującym).

Do pomiaru giętkości poznawczej stosuje się także test fluencji słownej polegający na wygenerowaniu w jednostce czasu jak największej liczby wyrazów zaczynających się na określonej literę (głoskę), a następnie należących do wspólnej kategorii semantycznej, np. zwierząt (Diamond, 2013). Zadanie takie wymaga przełączania uwagi (szczególnie jeśli oba kryteria stosowane są naprzemiennie a nie jedno po drugim) poprzez spontaniczny dobór słów (planowanie i generowanie) zakodowanych w leksykonie umysłowym zgodnie z podanym kryterium formalnym lub treściowym (Woźniak-Prus i Gambin, 2011; Jodzio i in., 2012). Tym samym zadanie fluencji słownej spełnia kryteria zadania, badającego elastyczność poznawczą (Józefacka-Szram, 2014), choć jednocześnie wykazuje brak lub niski poziom korelacji z Testem Sortowania Kart z Wisconsin (Szepietowska, 2010)<sup>85</sup>.

Kontaminacja pomiarów jednej funkcji przez zaangażowanie w wykonanie zadania pomiarowego innej funkcji jest znanym problemem badań nad funkcjami wykonawczymi. Według Diamond (2013) dzieje się tak, gdyż poszczególne komponenty funkcji wykonawczych są ze sobą ściśle splecione. Pamięć robocza wspiera kontrolę hamowania, odpowiadając za podtrzymanie dostępu do reguły wyznaczającej właściwy sposób zachowania. Z kolei hamowanie wspiera pamięć roboczą pozwalając oprzeć się wewnętrznym i zewnętrznym dystraktorom. Zarówno pamięć robocza, jak i hamowanie są ważne także dla elastyczności poznawczej, co zostało już omówione na przykładzie zadań przełączania. Wszystkie składowe dopełniają się więc we wspólnej funkcji kontroli zachowania, gotowości do działania ukierunkowanego na realizację celu, planowania oraz inicjowania aktywności (Rajtar i in., 2014).

Trudność w przeprowadzeniu izolowanego pomiaru poszczególnych funkcji ma przełożenie na badanie rozwoju funkcji wykonawczych w dzieciństwie. Skoro nie można do końca oddzielić w zadaniach hamowania od pamięci roboczej, czy przełączania od hamowania, to trudno jest śledzić trajektorie rozwojowe każdej z tych funkcji. Pewnym rozwiązaniem jest pomiar zmiennych latentnych (Lehto i in., 2003). Między innymi, dzięki

<sup>85</sup> Fluencja słowna jest także czasem traktowana, jako osobna funkcja wykonawcza, utożsamiana z testem służącym do jej pomiaru. W takim ujęciu jest to zdolność do tworzenia, planowania i wykonywania sekwencji niezautomatyzowanych czynności (Jodzio, 2008). Test fluencji słownej traktowany jest również, jako zadanie wyłącznie angażujące funkcje językowe (por. Króliczak i in., 2021).

nim wiadomo, że rozwój funkcji wykonawczych wraz z wiekiem jest nieharmonijny i postępuje stopniowo, osiągając pełnię sprawności we wczesnej dorosłości, co podyktowane jest zarówno dojrzewaniem struktur mózgowych (szczególnie obszarów czołowych) oraz wpływami środowiska fizycznego i społeczno-kulturowego (Brzezińska i Nowotnik, 2012). Poszczególne komponenty konstrukt „funkcje wykonawcze” różnią się trajektoriami swojego rozwoju i wraz z wiekiem coraz bardziej się różnicują (Putko, 2008; Best i Miller, 2010). Najwcześniejsze przejawy samoregulacji obserwuje się już u niemowląt. Intensywny rozwój funkcji takich jak hamowanie, przełączanie, planowanie, czy pamięć robocza następuje najpierw w ciągu pierwszych 18 miesięcy po urodzeniu (Woźniak-Prus i Gambin, 2011). Przy czym elastyczność poznawcza wykształca się najwolniej (Diamond, 2013). Trzyletnie dziecko potrafi na przykład włożyć w odpowiednie miejsce klocki różniące się kolorami, co wymaga przekierowywania uwagi dla dokonania wyboru (elastyczność poznawcza) i hamowania reakcji konkurencyjnych (Józefacka-Szram, 2014). Właśnie w wieku przedszkolnym następuje u dzieci intensywny rozwój funkcji wykonawczych (Woźniak-Prus i Gambin, 2011). W okresie szkolnym poziom sprawności funkcji wykonawczych stopniowo zbliża się do poziomu osób dorosłych (Best i Miller, 2010; Józefacka-Szram, 2014).

### **3.3 Kontrola procesu mówienia a funkcje wykonawcze**

Przyjęcie założenia, o intencjonalnym i celowym charakterze działania realizowanego w procesie mówienia, pozwala na postawienie pytań o relację przebiegu produkcji mowy (zwłaszcza jej monitorowania i kontroli) z działaniem funkcji zarządczych. Szczególna uwaga kierowana jest tu na procesy kontroli wykonawczej, które mogą mieć wpływ na różne aspekty przebiegu procesu mówienia, np. wykrywanie i reparację błędów. Analiza wpływu funkcji wykonawczych na zorganizowany przebieg intencjonalnego działania pozwala przewidywać, że mogą one mieć udział także w kontroli procesu mówienia.

Wbrew temu, dominujące na początku w psycholingwistyce modułowe rozumienie procesów językowych ugruntowało traktowanie mówienia i rozumienia mowy jako względnie automatycznych i autonomicznych wobec innych procesów poznawczych (Meyer i in., 2007). Wpływową w językoznawstwie i psycholingwistyce teorią Chomsky’ego traktowała język (gramatykę) jako moduł umysłu funkcjonujący pomiędzy systemem pojęciowo-intencjonalnym, a narządami wytwarzania i odbioru mowy, ale generujący zdania niezależne od nich (Kurcz, 2005; Mecner, 2008). Ten sposób myślenia o procesach językowych znalazł

do pewnego stopnia odzwierciedlenie w modelach procesu mówienia (por. rozdział 2.). Choć Levelt (1989) nie nazywał poszczególnych komponentów opisanej przez siebie architektury produkcji mowy terminem „moduł”, to cechy, jakie im przypisywał, tj. automatyzm, hermetyczność informacyjna i poznawcza nieprzenikliwość, każą traktować je jak typowe, autonomiczne moduły, odrębne od wyższych, bardziej kontrolowanych procesów umysłowych (por. Meyer i in., 2007). Dotyczyło to przede wszystkim pracy formulatora (a także artykulatora), w którym proces formułowania (kodowania językowego) Levelt (1989) określał jako w dużej mierze automatyczny i niedostępny (*impenetrable*) dla kontroli wykonawczej. Zastrzeżenie to nie dotyczyło konceptualizatora, który odpowiadać ma za wysoce kontrolowane procesy, wymagające zasobów uwagi oraz angażujące kontrolę wykonawczą, m.in. na potrzeby zatrzymania artykulacji w razie wykrycia błędu. Wynika z tego, że Levelt (1989) już w pierwszym modelu dostrzegał teoretyczną możliwość wiązania kontroli procesu mówienia z innymi procesami kontrolnymi. Nie rozwijał jednak tego wątku w opisie produkcji mowy. Jak zauważają Meyer i in. (2007), dominujące wówczas założenie o braku związku komponentów języka z innymi procesami poznawczymi skutecznie odwracało uwagę badaczy od powiązań produkcji mowy z funkcjami wykonawczymi.

Współcześnie Hartsuiker i Moors (2017) proponują odejście od dychotomicznego podziału na procesy automatyczne i nieautomatyczne, na rzecz bardziej stopniowalnego, płynnego przechodzenia między wymaganiami kontroli. Zarówno w procesach konceptualizacyjnych mogą zachodzić procesy samoczynne, nie wymagające kontroli, tak i w dostępie leksykalnym, czy kodowaniu gramatycznym ujawnia się częściowa zależność od innych zasobów poznawczych. Z tej perspektywy żadnego językowego procesu nie można traktować jako wyłącznie automatycznego lub nieautomatycznego. Autorzy ci dodatkowo czynią jeszcze rozróżnienie na podstawowe i wtórne procesy językowe. Podczas gdy procesy podstawowe obejmują to, co zwykle się nazywa produkcją mowy, procesy wtórne pełnią wobec tych pierwszych funkcję nadzorczą i interweniują w razie niepowodzeń. Zdaniem Hartsuikera i Moorsa (2017) konsekwencją zacierania tego rozróżnienia jest przypisywanie procesom podstawowym cech automatyzacji, podczas gdy pewne cechy automatyzacji może zdradzać także proces wtórny. Jednocześnie to procesy wtórne w największym stopniu współdzielą zasoby z innymi procesami poznawczymi, przez co podatne są na ich ograniczenia. Właśnie te kontrolne aspekty mówienia są – zdaniem Nozari (2018) – współdzielone przez system produkcji mowy i inne systemy poznawcze. Innymi słowy ogólne mechanizmy (*domain-general mechanisms*) mogą działać na reprezentacjach specyficznych dla danej domeny (*domain-specific representations*), np. języka, tworząc wyspecjalizowane

pętle monitorowania i kontroli służące wykrywaniu błędów w mówieniu. Mówiąc o ogólnych mechanizmach chodzi o funkcje wykonawcze, które odpowiadają za kontrolę celowego działania w różnych domenach, także w sferze produkcji mowy.

Choć pojęcie monitorowania i kontroli mówienia w podejściu opartym na odbiorze i rozumieniu mowy (Levelt, 1989; Levelt i in., 1999; Roelofs, 2005; Roelofs, 2020) różni się w szczegółach od stanowiska koneksjonistycznego, reprezentowanego przez Nozari (Nozari i in., 2011; Nozari, 2020), także ono zakłada, że kontrola mówienia opiera się na funkcjach wykonawczych. Według Roelofsa (2020a) „monitor”, pełniący rolę „krytyka”, porównuje reprezentacje użyte w produkcji mowy z reprezentacjami aktywowanymi przez system odbioru mowy. W przypadku wykrycia pomyłki sygnał błędu wysyłany jest do „kontrolera” (rola ta przypisywana jest funkcjom kory przedniego zakrętu obręczy), który modyfikuje wykonanie planu przez system produkcji mowy. Porównywalny mechanizm opisywany jest w teorii systemów, kiedy to monitorowanie, jako „słabsza” postać kontroli, nie wywołująca bezpośrednich zmian w nadzorowanym systemie, może wzbudzić silniejszą postać kontroli, czyli sterowanie lub regulację (Chuderski, 2010). Sterowanie oznacza sprawowanie pełnej kontroli jednego systemu przez drugi, natomiast regulacja polega na dokonywaniu zmian w kontrolowanym systemie w zależności od jego stanu. Stan systemu produkcji mowy określany jest przez stopień realizacji intencji komunikacyjnej (Levelt, 1989; por. Frydrychowicz, 1999) i zgodność reprezentacji w pętli „mówienie-rozumienie” (Levelt i in., 1999; Roelofs, 2020a) lub poziom konfliktu (Nozari i in., 2011; Nozari, 2020). Dzięki ciągłemu monitorowaniu tego stanu proces mówienia podlega bieżącym regulacjom, w które mogą być zaangażowane funkcje wykonawcze, szczególnie w sytuacjach zakłócenia przebiegu tego procesu, np. w wyniku wystąpienia pomyłki (lub ryzyka błędu spowodowanego dystrakcją). W dalszej części rozdziału przedstawione zostaną empiryczne przesłanki do tego, aby uznać, że w kontrolę procesu mówienia mogą być włączone funkcje wykonawcze<sup>86</sup>.

---

86 W tym miejscu warto odnotować, że w literaturze przedmiotu związku między funkcjami wykonawczymi a przetwarzaniem języka analizowane są dwukierunkowo (Jodzio, 2008; ). Z jednej strony przypisuje się językowi duże znaczenie w samokontroli zachowania człowieka, a z drugiej strony podkreśla się rolę jaką funkcje wykonawcze mogą pełnić w sprawnym posługiwaniu się językiem (Rajtar i in., 2014). Ten pierwszy kierunek zależności nie będzie szerzej omawiany, gdyż nie jest on przedmiotem niniejszej pracy, ponadto brakuje konsensusu dotyczącego tego w jaki sposób mowa lub język miałyby przyczyniać się do regulacji zachowania. Kluczowym pojęciem jest „mowa wewnętrzna”, wprowadzona przez Wygotskiego i stosowana przez Łurję (por. Maruszewski, 1970). Brakuje jednak teoretycznych podstaw aby łączyć „mowę wewnętrzną” (*inner speech*) służącą organizacji zachowania z „mową wewnętrzną” (*internal speech*) w rozumieniu, jakie nadaje temu terminowi psycholingwistyka (por. Levelt, 1989; Levelt i in. 1999), choć należy odnotować pewne próby w tym zakresie poczynione przez Gevę i Fernyhough (2019). W związku z tym, że przedmiotem niniejszej rozprawy są poznawczo-rozwojowe aspekty kontroli mówienia, w dalszej części rozdziału omawiany będzie jedynie drugi kierunek zależności, rozpatrujący znaczenie elementarnych funkcji wykonawczych dla kontroli produkcji mowy.

### 3.3.1 Zadania podwójne (*dual tasks*)

W codziennych działaniach, wykonując jedną czynność często myślimy już o innej, którą będziemy wykonywać za chwilę, a może dopiero planujemy jakieś działania w przyszłości. Często też próbujemy pogodzić realizację kilku zadań jednocześnie, np. przygotowując posiłek lub wykonując inne rutynowe działania. W takich sytuacjach musimy dzielić zasoby uwagowe na różne czynności, co wymaga zaangażowania funkcji wykonawczych. Nietrudno wtedy o pomyłkę spowodowaną przekroczeniem możliwości kontroli wykonawczej. Reason (1984) podaje różne przykłady takich pomyłek (*lapses of attention*), np. zapomnienie z jaką intencją rozpoczęło się daną czynność, zrobienie czegoś wbrew pierwotnej intencji, szukanie obiektu który ma się w ręce, wykonanie właściwej czynności, ale na niewłaściwych przedmiotach lub w odwróconej kolejności, sięgnięcie po niewłaściwy obiekt, powtórzenie czynności, pojawienie się intruzji z innej czynności itp.. Również w mówieniu obserwujemy podobne problemy i pomyłki: zapomnienie, o czym chciało się powiedzieć, efekt „mam na końcu języka”, przejęzyczenia spowodowane myśleniem o czymś innym. Na przykład ktoś odczuwający głód i planujący żeby coś zjeść może powiedzieć „Zjadłem wszystkie moje biblioteczne książki” zamiast „Przeczytałem wszystkie moje biblioteczne książki” (Harley, 2001; por. Harley, 1984)<sup>87</sup>.

Naturalną sytuację konieczności rozdysponowania ograniczonych zasobów poznawczych podczas wykonywania dwóch czynności jednocześnie symulują tzw. zadania podwójne (*dual tasks*) (Piotrowski, Stettner, Wierzchoń, Balas i Bielecki, 2009). Przykładem takiego zadania jest wariant monitorowania fonemów opisany przez Wheeldon i Levelta (1995). Przypomnijmy, osoby badane podzielone na dwie grupy miały za zadanie monitorować tłumaczone z obcego języka wyrazy w poszukiwaniu określonych fonemów (zadanie główne). Dodatkowo jedna z tych grup otrzymała zadanie odliczania na głos (zadanie doładujące), co miało powodować supresję pętli artykulacyjnej (składnika pamięci roboczej, por. Baddeley, 1998). Badaczom chodziło o to, żeby sprawdzić, czy reprezentacja dostępna monitorowaniu mowy ma charakter fonologiczny, czy fonetyczny lub artykulacyjny (por. Levelt, 1989; Levelt, 1999). Dodatkowe zadanie wprowadziło istotne statystycznie, choć niewielkie odchylenia w poprawności wykonania zadania oraz spowodowało jego opóźnienie, względem średniej w grupie osób badanych bez zadania odliczania (dłuższy czas reakcji w udzielaniu odpowiedzi). Wynik ten można więc uznać za dowód na to, że stopień obciążenia pamięci roboczej (odliczanie angażuje pętlę

---

<sup>87</sup> Org. Target: *I've read all my library books.* Utterance: *I've eaten all my library books.*  
Context: *The speaker reported that he was hungry and was thinking of getting something to eat.* (Harley, 2001, s. 380).

artykulacyjną) ma znaczenie dla efektywności monitorowania procesu mówienia. Wheeldon i Levelt (1995) nie analizowali jednak tego wątku, koncentrując się na tym, że niewielkie różnice między wynikami obu grup (oraz identyczny w obu grupach wzorzec odpowiedzi) są potwierdzeniem na fonologiczną naturę reprezentacji dostępnej monitorowaniu. Można jednak przewidywać, że różnice indywidualne w funkcjonowaniu pamięci roboczej mogą przekładać się na efektywność monitorowania procesu mówienia.

Korzystanie ze wspólnych zasobów przez kontrolę procesu mówienia i funkcje wykonawcze sugerują też wyniki przedstawione przez Oomen i Postmę (2002, za: Postma i Oomen, 2005). Zadanie podwójne, które postawili przed osobami badanymi polegało na opisywaniu obrazków (*picture-description task*) przy jednoczesnym losowym stukaniu palcami (*random finger tapping*)<sup>88</sup>. W porównaniu do warunków kontrolnych, dodatkowe zadanie wiązało się ze zwiększeniem liczby popełnianych błędów fonologicznych, przy jednoczesnym spadku liczby reperacji błędów w mówieniu. Widać więc, że kontrola nad procesem mówienia, zarówno w zakresie programowania mowy, jak i późniejszej ewaluacji pogarsza się, kiedy równoległe z mówieniem wykonywana jest jeszcze jedna, wymagająca kontroli wykonawczej, nieautomatyzowana czynność (por. Hartsuiker, 2007).

Nawet w przypadku czynności złożonych, ale względnie poddających się automatyzacji, jak prowadzenie pojazdu, zasoby poznawcze jakie przeznaczone są na to działanie i na mówienie mogą być ograniczone. Bock, Dell, Garnsey, Kramer, i Kubose (2007) badali jak słuchanie lub mówienie wpływa na stabilność prowadzenia samochodu (prędkość, utrzymanie pasa ruchu, utrzymanie odstępu) w symulatorze poruszania się w wirtualnym otoczeniu. Zgodnie z przewidywaniami dodatkowe zadanie pogarszało<sup>89</sup> wykonanie zadania głównego, choć nie było różnicy między efektem mówienia a efektem słuchania.

---

88 Zadanie stukania palcami polega na wystukiwaniu palcami niedominującej dłoni sekwencji cyfr 1-4, które przypisane są poszczególnym palcom.

89 Warto przy tym odnotować, że wbrew przewidywaniom i poniekąd paradoksalnie (szczególnie względem innych wyników) utrzymanie pozycji na pasie ruchu wykazywało się mniejszą zmiennością podczas mówienia, niż kiedy zadanie było wykonywane w ciszy. Być może więc, w zadaniach, które do pewnego stopnia są zautomatyzowane, przesunięcie „uwagi” monitora na inne domeny (np. mówienie), może sprzyjać zwiększeniu płynności w wykonywaniu czynności związanych z zadaniem. Pamiętać jednak należy, że pozostałe parametry panowania nad pojazdem pogorszyły się w wyniku zaangażowania osób badanych w dodatkowe zadanie mówienia lub słuchania.

### 3.3.2 Dane kliniczne

Coraz częściej podkreśla się związek funkcji wykonawczych z mówieniem w kontekście obrazu klinicznego objawów prezentowanych przez pacjentów z diagnozą różnych postaci afazji. Jodzio (2008) zauważa, że w wyniku uszkodzenia części przedczołowej lewego płata czołowego może wystąpić tzw. afazja dynamiczna (inna nazwa to transkorowa afazja ruchowa), objawiająca się problemami z tworzeniem mowy spontanicznej, przy zachowaniu funkcji rozumienia i powtarzania. Jego zdaniem w tej postaci afazji główną przyczyną utrudnień w wyrażaniu myśli w mowie jest specyficzny deficyt funkcji wykonawczych, ograniczający przechodzenie od zamiaru (sformułowanie intencji) do wykonania (wyartykułowanie wypowiedzi językowej).

Coraz liczniejsze prace wykazują, że chorzy z diagnozą różnych postaci afazji gorzej wykonują także testy mierzące poziom funkcji wykonawczych (Rajtar i in., 2014). Na przykład Nozari (2018) przywołuje przypadek pacjenta, który w zadaniu nazywania obrazków popełniał wiele błędów o podłożu semantycznym (np. podawał nazwy różnych owoców, kiedy obrazek przedstawiał arbuza). Osoba ta, jednocześnie wykazywała obniżony poziom kontroli hamowania w standardowych testach funkcji wykonawczych. Sugeruje to, że kontrola wykonawcza może odgrywać istotną rolę na etapie selekcji leksykalnej lub, że przynajmniej mechanizmy kontrolne selekcji leksykalnej funkcjonują na bazie obwodów neuronalnych, których funkcją jest także hamowanie ujawniane w testach diagnostycznych.

U innych pacjentów, opisywanych przez Humphreysa i in. (2007) trudności w nazywaniu (tzw. anomia) współwystępowały również z deficytami w zakresie funkcji wykonawczych. Z kolei Lambon Ralph i Fillingham (2007) zwracają uwagę na to, że osoby z afazją różniące się pod względem poziomu funkcji wykonawczych odnoszą różne korzyści z terapii. Badacze porównywali strategię opartą na uczeniu się na podstawie błędów (*errorful learning*) z metodami ograniczającymi popełnianie błędów przez tzw. uczenie się bezbłędne (*errorless learning*). Według nich osoby z najniższymi umiejętnościami uwagowo-wykonawczymi (*attentional-executive skills*) nie są w stanie wykryć popełnianych błędów, przez co terapia nie przynosi pożądanego rezultatu, w odróżnieniu od osób, które przejawiają wyższy poziom tych umiejętności i są w stanie uczyć się zarówno na błędach, jak i w strategii uczenia bezbłędnego.

Współwystępowanie zaburzeń kontroli wykonawczej i produkcji oraz kontroli mowy daje podstawy do przypuszczeń o wspólnych obwodach neuroanatomiczno-funkcjonalnych dla obu zdolności (Rajtar i in., 2014). Wsparciem dla klinicznych danych neuropsychologicznych są tu badania neuroobrazowania, które również dostarczają ważnego

wsparcia dla twierdzenia o częściowo wspólnym neuronalnym podłożu mówienia i funkcji wykonawczych.

### **3.3.3 Dane z badań metodami neuroobrazowania**

Mówienie jest złożoną czynnością obejmującą przetwarzanie języka na kilku poziomach. Nie powinno więc dziwić, że w czasie nawet tak stosunkowo prostej aktywności werbalnej, jaką jest generowanie pojedynczych wyrazów (np. w zadaniu nazywania obrazków) aktywne są różne obszary mózgu, przy czym kluczowe obserwacje dotyczą kory lewej półkuli mózgu<sup>90</sup>. Czasową charakterystykę tej aktywności (dopasowując ją do modelu Levelta i in., 1999) przedstawili Indefrey i Levelt (2004) w metaanalizie danych z badań metodami neuroobrazowania. Procesy kontrolne, związane z monitorowaniem mowy zewnętrznej angażują przede wszystkim obupółkulowo górny zakręt skroniowy (*superior temporal gyrus* – STG), który odpowiada za procesy odbioru i rozumienia mowy<sup>91</sup>. W późniejszych latach Indefrey (2011) aktualizował zestawienia obszarów mózgu zaangażowanych w generowanie wyrazów. Cytowane przez niego badania wykazują, że w monitorowanie własnej mowy zaangażowana jest szersza sieć neuronalna, obejmująca korę zakrętu obręczy, wyspę (*insula*), drugorzędowy obszar ruchowy, obustronne obszary ruchowe, mózdzek oraz struktury podkorowe (wzgórze i jądra podstawne). Część tej sieci jest więc współdzielona ze strukturami przypisywanymi funkcjom wykonawczym. Potwierdzają to także badania nie objęte metaanalizą.

Na przykład Xue i in. (2008) wykazali, że hamowanie reakcji manualnych i werbalnych w zadaniu sygnału stop angażuje te same obszary prawego dolnego zakrętu czołowego (*inferior frontal gyrus* – IFG) oraz kory przedruchowej (*pre-supplementary motor area* – pre-SMA). O ile inicjowanie reakcji werbalnej (wypowiadanie pseudosłów) specyficznie wiązało się z aktywacją tzw. obszaru Broca, czyli lewego dolnego zakrętu obręczy, o tyle powstrzymanie tej reakcji wykazywało podobne do hamowania reakcji manualnej (naciskanie klawisza) korelaty neuronalne. Istotną rolę prawego dolnego zakrętu czołowego w hamowaniu dystrakcji podczas mówienia wykazali także Severens i in. (2012). Badali oni tłumienie wyrazów dystrakcyjnych w zadaniu nazywania obrazków z interferencją.

---

90 Funkcje językowe są silnie zlateralizowane w ludzkim mózgu (Kroliczak i in., 2021). Levelt i Indefrey (2004) wyróżniają 11 obszarów lewopółkulowych oraz 4 obszary prawopółkulowe, tworzące neuronalną sieć produkcji wyrazów.

91 Obszar tylny lewego górnego zakrętu skroniowego (*left posterior superior temporal gyrus*) nazywany jest polem Wernickego. Jego uszkodzenie powoduje tzw. afazję czuciową (inaczej sensoryczną, akustyczną lub akustyczno-gnostyczną) objawiającą się znacznym upośledzeniem rozumienia tekstów słyszanych, także własnej mowy (Maruszewski, 1970).

Wyrazy tabu były skutecznie blokowane przez osoby badane, a zaangażowanie w ten proces prawego dolnego zakrętu czołowego było silniejsze niż podczas blokowania wyrazów neutralnych<sup>92</sup>.

Abel i in. (2009) wykazali, że obszary przedczołowe obu półkul wykazują silną aktywację podczas nazywania obrazków zarówno, kiedy osoba badana popełnia błąd, a także (z mniejszą intensywnością), kiedy prawidłowo generuje właściwy wyraz. Badacze interpretują ten wynik jako dowód na to, że monitorowanie procesu mówienia nie ogranicza się tylko wykrywania błędów, ale działa jako bardziej generalny mechanizm sprawdzający poprawność produkowanych wyrazów, przyczyniając się do ich korekty jeszcze przed wystąpieniem błędu. W tym samym badaniu zaobserwowano również, że w przypadkach udzielenia błędnych odpowiedzi silniej (niż w próbach bezbłędnych) aktywowane były obszary kory przedniego zakrętu obręczy (ACC), okolice pierwszo- i drugorzędowej kory ruchowej oraz kory przedruchowej, struktury podkorowe (jądra podstawne i wzgórze), a także – w mniejszym stopniu – prawa półkula mózdzku oraz górny zakręt skroniowy.

Aktywację podobnej sieci neuronalnej zaobserwowali Gauvin i in. (2016) w badaniu, którego uczestnicy mieli za zadanie wypowiadać łamańce językowe (*tongue twisters*). Badacze raportowali, związaną z wykrywaniem błędów, wzmożoną aktywność w obrębie kory przedniego zakrętu obręczy, obszarów przedczołowych i przedruchowych, pierwszo- (zakręt przedśrodkowy) i drugorzędowej kory ruchowej oraz struktur podkorowych i (obustronnie) kory wyspy. Ponownie więc wykazano, że neuronalne substraty monitorowania i (ogólnie) kontroli produkcji mowy pokrywają się w dużym stopniu z obszarami odpowiedzialnymi za kontrolę wykonawczą (por. Nozari, 2018).

### **3.3.4 Badania korelacyjne**

Choć niewiele jest publikacji na temat różnic indywidualnych w kontroli procesu mówienia, wiążących zmienność międzyosobową wśród badanych grup ze sprawnością funkcji wykonawczych, można wskazać kilka badań, które pokazują wzajemne relacje między tymi konstruktami. Na przykład van den Wildenberg i Christoffels (2010) odnotowali wysoką korelację czasów reakcji między odpowiedziami udzielanymi werbalnie i manualnie w wykonaniu zadania sygnału stop. Co więcej, badacze wykazali także, że osoby charakteryzujące się „dysfunkcyjną impulsywnością” (tendencja do działania z niewielką

<sup>92</sup> Co ciekawe, badacze (Severens i in., 2012) odnoszą swoje wyniki do populacji klinicznej osób z zespołem Touretta, u których obserwuje się trudne do zahamowania tiki ruchowe i wokalne. Powołują się przy tym na badania, które sugerują, że zredukowanie istoty białej w prawym dolnym zakręcie czołowym może być neuroanatomicznym podłożem tego zaburzenia.

przezornością) mają większe problemy z wyhamowaniem reakcji werbalnych niż osoby cechujące się bardziej adaptacyjną formą szybkiego reagowania („funkcjonalna impulsywność”).

Wardlow (2013) sprawdzała z kolei, czy opisywanie obrazków, uwzględniające perspektywę rozmówcy może zależeć od pamięci roboczej i funkcji hamowania. Zadanie przed jakim zostały postawione osoby badane wymagało nazywania obiektów widocznych na ekranie poprzez podanie nazwy obiektu (np. koło), a także – jeśli z perspektywy rozmówcy było to istotne – cechy odróżniającej wskazywany obiekt od innego obiektu z zestawu (dla odróżnienia koła od trójkąta wystarczy podać nazwę wskazywanej figury geometrycznej, ale dla odróżnienia dwóch trójkątów konieczne jest dodanie różnicującej je cechy, np. „mały trójkąt”). Przeprowadzone badanie wykazało, że prezentowany przez uczestników poziom funkcji wykonawczych koreluje z adekwatnością podawanych przez nich informacji w zadaniu nazywania. Na przykład im niższe wyniki badani uzyskali w zakresie pamięci roboczej, tym częściej tworzyli odniesienia nieskorygowane perspektywicznie. Z kolei wyższe wyniki w zadaniu kontroli hamowania wiązały się z rzadszym używaniem nieistotnych – z punktu widzenia rozmówcy – informacji na temat opisywanych obiektów.

Inne badanie przeprowadzone przez Shao i in. (2012) wykazało, że czas reakcji werbalnej w zadaniu nazywania obrazków istotnie negatywnie koreluje z miarą odświeżania pamięci roboczej, kiedy nazywane są czynności a nie przedmioty (szybkość nazywania przedmiotów nie korelowała istotnie z miarą pamięci roboczej). Podobne związki badacze wykazali w przypadku hamowania i szybkości nazywania obrazków. Czas reakcji w nazywaniu przedmiotów i czynności istotnie korelował z czasem reakcji w zadaniu sygnału stop rozwiązywanym manualnie, natomiast nie wykazywał istotnych związków z miarą giętkości poznawczej (zadanie przełączania między kategorią kształtu i koloru, *shape-colour switching task*). Ci sami badacze Shao i in. (2013) wykazali również korelację między kontrolą hamowania a czasem reakcji werbalnej w zadaniu nazywania obrazków z dystraktorem wywołującym interferencję (por. Nozari, 2018)<sup>93</sup>.

---

93 W dokładniejszej analizie, jako miarę hamowania selektywnego (*selective inhibition*) Shao i in. (2013) uwzględnili jedynie wyniki z ostatniego kwantyla czasów reakcji w nazywaniu obrazków (segment najwolniejszych czasów reakcji). Tak opracowany wskaźnik hamowania selektywnego nie wykazywał istotnych korelacji z hamowaniem nieselektywnym (*nonselective inhibition*), mierzonym przez zadanie sygnału stop. Według autorów jest to dowód na pewną rozdzielność między tymi dwoma rodzajami hamowania.

### 3.4 Poznawcza kontrola procesu mówienia

Podsumowując dotychczasowe ustalenia, przedstawione w niniejszym rozdziale psycholingwistyczne modele kontroli procesu mówienia koncentrują się głównie na ludzkiej zdolności do samodzielnego monitorowania na bieżąco produkowanej mowy. Panuje ogólny konsensus dotyczący tego, że oprócz monitorowania zewnętrznego, opartego głównie na kanale słuchowym, istnieje także monitorowanie wewnętrzne, umożliwiające wykrycie potencjalnych błędów jeszcze przed ich wyartykułowaniem. Najbardziej wpływowa w omawianym nurcie badań teoria pętli percepcyjnej (Levelt, 1989; Levelt i in., 1999) przyjmuje założenie o jednorodnej naturze monitorowania zewnętrznego i wewnętrznego. Zgodnie z tym założeniem ludzie monitorują swoje wypowiedzi za pośrednictwem systemu odbioru i rozumienia mowy, który ma dostęp zarówno do wypowiedzi wyartykułowanej (pętla zewnętrzna monitorowania korzysta z kanału słuchowego), jak i do mowy wewnętrznej (pętla wewnętrzna monitorowania korzysta z fonetycznej i/lub fonologicznej reprezentacji powstającej przed etapem artykulacji). Chociaż współczesne, konkurencyjne propozycje zgodne są co do roli pętli zewnętrznej, opartej na odbiorze mowy, to brak jest konsensusu w sprawie natury wewnętrznego monitorowania (por. Lind i Hartsuiker, 2020). Niektórzy badacze twierdzą, że monitorowanie wewnętrzne nie korzysta z systemu odbioru mowy, a jedyne informacje potrzebne do wykrywania błędów dostarcza sam system produkcji mowy (Nozari i in., 2011; Nozari, 2018; 2020). Mimo dostarczanych argumentów na rzecz tej hipotezy, teoria monitorowania oparta na odbiorze mowy nie została sfalsyfikowana i nadal użytecznie wyjaśnia różne zjawiska dotyczące kontroli mówienia, radząc sobie również z tymi, które przytaczane są jako argumenty przeciw jej trafności (Roelofs, 2020a; 2020b). Pomimo różnic dotyczących fundamentalnych kwestii, mających swoje źródła w odmiennym modelowaniu przebiegu produkcji mowy, badacze zgadzają się, że monitorowanie mowy wymaga zasobów poznawczych, których przydział organizowany jest przez funkcje wykonawcze. Przytaczane badania pokazują, że kontrola procesu mówienia oraz kontrola wykonawcza „krzyżują” się zarówno pod względem funkcjonalnym, jak i na poziomie neuroanatomicznym.

Przyjmując, na bazie opisanych badań empirycznych, że kontrola produkcji mowy wykazuje związki z funkcjami wykonawczymi, których poziom do pewnego stopnia przewiduje jej sprawność, można postawić pytanie o to, na ile procesy dojrzewania układu nerwowego i nabywania doświadczenia wraz z wiekiem, odzwierciedlające się w rozwoju funkcji wykonawczych, znajdują swoje odbicie także w zmianach kontroli procesu mówienia.

Problem rozwoju umiejętności sprawowania kontroli nad procesem mówienia jest przedmiotem kolejnego rozdziału.

## 4 Rozwojowe aspekty kontroli procesu mówienia

W poprzednich dwóch rozdziałach przedstawione zostało psycholingwistyczne podejście do produkcji mowy oraz jej kontroli przez wewnętrzną i zewnętrzną pętlę monitorowania. Dzięki licznym badaniom w tym obszarze wiele już dowiedziano się na temat poznawczej kontroli procesu mówienia. Nie można jednak tego samego powiedzieć o badaniach nad rozwojem kontroli mówienia. Do tej pory w tej kwestii przeprowadzono stosunkowo niewiele badań, szczególnie jeśli chodzi o badania w warunkach kontrolowanych. Większość wiedzy na temat zdolności do monitorowania procesu mówienia u dzieci pochodzi z mniej lub bardziej ustrukturyowanych obserwacji w domu lub klasie (Hanley i in., 2016). W niniejszym rozdziale podjęta zostanie próba nakreślenia tego, jak zmienia się wraz z wiekiem kontrola procesu mówienia na podstawie przeprowadzonych do tej pory badań.

Żeby jednak rozważać kwestię poznawczej kontroli mówienia, trzeba najpierw odpowiedzieć na pytanie, w którym momencie rozwoju dzieci zaczynają mówić, to znaczy kiedy w rozwoju komunikacji i nabywania języka można powiedzieć o „mówieniu”, jako o złożonym procesie transformacji pojęć i intencji w słowa wyrażane w postaci artykułowanych dźwięków mowy (zgodnie z modelami przedstawionymi w rozdziale 2.). Kwestia rozwoju mówienia zostanie więc rozpatrzona przed opisaniem badań na temat monitorowania i kontroli mówienia u osób w różnym wieku.

### 4.1 Rozwój procesu mówienia

Badania nad rozwojem mowy i nabywaniem języka mają swoją długą tradycję w psychologii. Dla dyscypliny psycholingwistyki jest to w zasadzie jeden z trzech filarów określających kierunek podejmowanych badań (Bock, 1996; Dell, 1986). Choć wiele uwagi poświęcono na badanie rozwoju słownictwa, gramatyki, fonologii i wielu innych aspektów języka i komunikacji, badania rozwojowe nad mówieniem i kontrolą mówienia należą do rzadkości. Przykładowo, Levy (1999) analizując przypadki dziecięcych korekt w ich własnych wypowiedziach opiera się na modelu Levelta (1989) – opracowanym na danych dostarczonych przez dorosłych użytkowników języka – sugerując jednocześnie możliwe różnice w sposobie przebiegu procesu produkcji mowy u dzieci i dorosłych. Jak zauważa Rispoli (2003, s. 821) interpretacja części badań dokonuje się przez „urozwojowienie” (*developmentalization*) modelu dotyczącego przebiegu mówienia u osób dorosłych.

W niniejszym podrozdziale, w oparciu o uwagi samego Levelta (1998) i oryginalne badania nad mową dzieci, przeprowadzona zostanie rekonstrukcja tego, w jaki sposób tworzą one wypowiedzi językowe i jak bardzo proces ten komplikuje się w pierwszych latach życia, dochodząc do dojrzałej lub prawie dojrzałej postaci, znanej z badań nad osobami dorosłymi. Użycie słowa „rekonstrukcja” oznacza zarysowanie pewnego możliwego trendu w rozwoju mówienia na podstawie dostępnych danych, gdyż jak dotąd – według wiedzy autora – nie przeprowadzono systematycznego badania longitudinalnego, pozwalającego na opracowanie modelu rozwojowego procesu mówienia<sup>94</sup>.

#### **4.1.1 Ontogenetyczne źródła dualnej struktury procesu mówienia**

Levelt i in. (1999) wprowadzając swój model dostępu leksykalnego podczas generowania słów zauważyli, że w produkcji mowy osób dorosłych można wyróżnić dwa stosunkowo odrębne systemy, tj. system retoryczno-semantyczno-syntaktyczny (pojęciowy) i fonologiczno-fonetyczny (językowy), które wiążą się z rozdzieleniem reprezentacji na poziomie słownika lemm i słownika słowo-form (morfemów). Ten rozdział widoczny jest na przykład w zjawiskach (a) „mam na końcu języka” (*tip-of-the-tongue*), ale też w (b) różnej dystrybucji przejęczyzeń semantycznych i fonologicznych oraz w (c) odmiennych kierunkach wpływu dystrakcji semantycznej i fonologicznej na nazywanie obrazków (Schriefers i in., 1990): semantyczne związki między dystrakcyjnym wyrazem, a nazywanym obrazkiem z reguły dają efekt interferencji semantycznej i opóźniają podawanie nazwy obrazka, a związki fonologiczne z reguły przyspieszają ten proces, co nazywane jest facylitacją fonologiczną (efekt interferencji i facylitacji jest modyfikowany relacją czasową między prezentacją obrazka a podaniem dystraktora). Według autorów modelu (Levelt i in., 1999) rozdział między systemem retoryczno-semantyczno-syntaktycznym i fonologiczno-fonetycznym ma swoje źródło w ontogenezie.

Nie jest to zupełnie nowa myśl w kwestii rozumienia, w jaki sposób człowiek rozwija typowo ludzką zdolność posługiwania się językiem w mowie. Jednym z pierwszych – jeśli nie pierwszym – badaczem rozwoju, który zwrócił na to uwagę był Wygotski (1989/1934). Twierdził on, że mowa i myślenie mają różne źródła genetyczne zarówno w filogenezie, jak i w ontogenezie, tzn. można wyróżnić „fazę przedjęzykową” w rozwoju myślenia oraz fazę „przedintelektualną” w rozwoju mowy. Zostawiając na boku kwestie ewolucji i powstania

---

94 Jak dotąd przeprowadzono wiele badań longitudinalnych i poprzecznych nad nabywaniem języka, które cytowane są w niniejszym rozdziale, jednak skoncentrowane są one zazwyczaj na raportowaniu postępów w rozwoju jakiegoś aspektu systemu językowego niż na zmianach rozwojowych w produkowaniu mowy. Stąd wiedza o tym, jak rozwija się proces mówienia ma charakter fragmentaryczny.

języka (filogeneza), Wygotski (1989/1934, s. 51) uważa, że do pewnego momentu w rozwoju dziecka (ontogeneza) „linie rozwoju myślenia i mowy” przebiegają oddzielnie, krzyżując się w drugim roku życia. To, co dla innych badaczy było nagłym „odkryciem” języka, dla Wygotskiego stanowi efekt połączenia dwóch niezależnie rozwijających się ścieżek myślenia i mowy (mowa rozumiana jest przez Wygotskiego jako komunikowanie się za pomocą dźwięków).

Powołując się na badania Köhlera, Wygotski (1989/1934) stwierdza, że w przedjęzykowym stadium myślenia ma ono charakter instrumentalny, to znaczy, że czynności stają się świadomie usensownione, a więc świadomie celowe. Innymi słowy, osiągnięta na tym poziomie inteligencja ma charakter praktyczny, gdyż rozwijana jest zdolność działania celowego, w pełni intencjonalnego (por. Piaget, 1966). Równolegle, zdaniem Wygotskiego rozwija się dźwiękowy aspekt komunikacji (mowa), który znajduje się w pierwszym roku życia w stadium przedintelektualnym. Linia rozwoju mowy na tym etapie przebiega od krzyku, przez gaworzenie do pierwszych słów, które - według Wygotskiego - nie mają jeszcze charakteru w pełni znakowego, gdyż funkcjonują bardziej jako bodźce (lub reakcje) warunkowe.

Sytuacja zmienia się, kiedy „w drugim roku życia (...) linie rozwoju myślenia i mowy, przebiegające dotąd oddzielnie, zaczynają się krzyżować, pokrywać i dają początek zupełnie nowej, tak charakterystycznej dla człowieka formie zachowania się” (Wygotski, 1989/1934, s. 51). Wygotski odnotowuje, że po tym momencie dziecko zaczyna aktywnie wzbogacać leksykon umysłowy, powiększając zasób słownictwa. Dzięki przecięciu się obu dróg rozwoju mowa stopniowo zaczyna się intelektualizować, a myślenie – werbalizować, to znaczy być „myśleniem opartym na słowie” (Wygotski, 1989/1934, s. 53). Według Wygotskiego (1989/1934, s. 51-52) dziecko „odczuwa potrzebę słów i dąży aktywnie do opanowania znaku, który odpowiada przedmiotowi, znaku który służy do nazywania i komunikowania”.

Idąc tym semiotycznie zorientowanym tropem, można powiedzieć, że w znakach języka skupia się wspólna, skrzyżowana droga rozwoju mowy i myślenia. Jednocześnie Wygotski zastrzega, że to skrzyżowanie się, czy przecięcie obu dróg nie musi być jednorazowe, ale że może przebiegać etapami i występować wielokrotnie, obie linie mogą przecinać się w wielu punktach i narastać stopniowo, rozciągając cały proces aż po wiek szkolny (Wygotski, 1989/1934).

Próbując odnieść ten rozwojowy proces do teorii językoznawczej, można porównać przecięcie się obu – wyróżnionych przez Wygotskiego (1989/1934) dróg rozwoju – do połączenia struktur, czy też reprezentacji semantycznych ze strukturami (reprezentacjami)

fonologicznymi. Są to niezbędne elementy znaku językowego odnoszące się do jego strony pojęciowej i dźwiękowej, które łączy relacja symbolizacji (Taylor, 2007). W mówieniu obie drogi ulegają powiązaniu, gdyż wypowiedź (znak językowy), np. pojedyncze słowo łączy w sobie struktury semantyczne (znaczeniowe) i dźwiękowe. Co istotne, połączenie obu struktur w znaku językowym podczas mówienia, powoduje, że oddziałują one wzajemnie na siebie. Jak pisał Wygotski (1989/1934, s. 51) „mowa stopniowo intelektualizuje się, a myślenie werbalizuje”, co sugeruje że zarówno na poziomie struktur semantycznych, jak i fonologicznych w wyniku procesów rozwojowych będzie dochodziło do przekształceń, modyfikujących w efekcie także sam proces mówienia.

Dla zilustrowania tych zmian warto w tym miejscu przypomnieć model znaku językowego przedstawiony przez Hjelmsleva (1979, zob. rozdział 2. niniejszej rozprawy). Oczywiście nie można bezpośrednio zestawić ze sobą teorii Wygotskiego ani Hjelmsleva, ale mogą one stanowić pewien punkt odniesienia dla propozycji rozwoju produkcji mowy według Levelta (1998), która zostanie przedstawiona w dalszej części tego podrozdziału.

Według Hjelmsleva (1979) znak językowy składa się z planu treści oraz planu wyrażania, które odpowiadają temu, co językoznawstwo kognitywne nazywa reprezentacjami (strukturami): semantyczną i fonologiczną. Zarówno plan treści, jak i plan wyrażania składają się ze swojej substancji i formy. Jak zauważono w rozdziale 2. w procesie mówienia (w każdym akcie mowy) dochodzi do dynamicznego powiązania substancji treści (tego, co może zostać pomyślane pojęciowo) z substancją planu wyrażania (dźwiękami języka, które można wyartykułować) za pomocą odpowiedniej formy. Ma to miejsce zarówno w wypowiedzianiu pierwszych słów przez bardzo małe dzieci, kiedy – jak to ujmuje Wygotski – linie rozwoju mowy i myślenia dopiero zaczynają się przecinać, jak i w rozbudowanej wypowiedzi osoby dorosłej, której mowa jest już w pełni „intelektualna”, a myślenie ma charakter „werbalny”. Obie sytuacje różni jednak to, w jaki sposób substancja planu treści i planu wyrażania wiążą się ze sobą. W wypowiedzi osoby dorosłej jest wyraźnie zarysowana forma językowa, którą jeszcze trudno dostrzec w dziecięcych, jednowyrazowych wyrażeniach. Jak zauważa Frydrychowicz (1988), początkowo związek między planem treści a planem wyrażania ma charakter przypadkowy, a dziecko odnosi generowane dźwięki bezpośrednio do przedmiotu odniesienia. Niejednokrotnie więc stosuje te same sylaby (jedna forma dźwiękowa) na określenie różnych obiektów i zdarzeń lub – na odwrót – używa różnych wyrażen na oznaczenie tego samego elementu rzeczywistości pozajęzykowej (por. Łobacz, 2005).

Jeśli jednak za Wygotskim przyjąć, że relacja między mową i myśleniem nie jest stała (zob. płaszczyzny mówienia opisane w rozdziale 2.) oraz, że w toku rozwoju dziecka

wpływają one na siebie, wzajemnie się organizując („mowa stopniowo intelektualizuje się, a myślenie werbalizuje”), to możliwe jest też przyjęcie, że w pierwszych dziecięcych słowach substancja treści i wyrażania mniej lub bardziej bezpośrednio łączą się ze sobą, z czasem dopiero ulegając formalizacji. Inaczej mówiąc proces rozwojowy obserwowany po 1 r. ż. dziecka, kiedy podejmuje ono próby intencjonalnego mówienia (generowania znaków), mógłby polegać na przekształcaniu planów treści i wyrażania poprzez nadawanie im formalnej organizacji dzięki rozwojowi systemu fonologicznego oraz przyswajaniu reguł gramatycznych. Dzięki tej formalnej specyfikacji połączeń planu treści i wyrażania ich – początkowo przypadkowy i umotywowany podmiotową aktywnością dziecka – związek może stać się w pełni arbitralny (Frydrychowicz, 1988), a używane przez dziecko słowa – w pełni znakami języka.

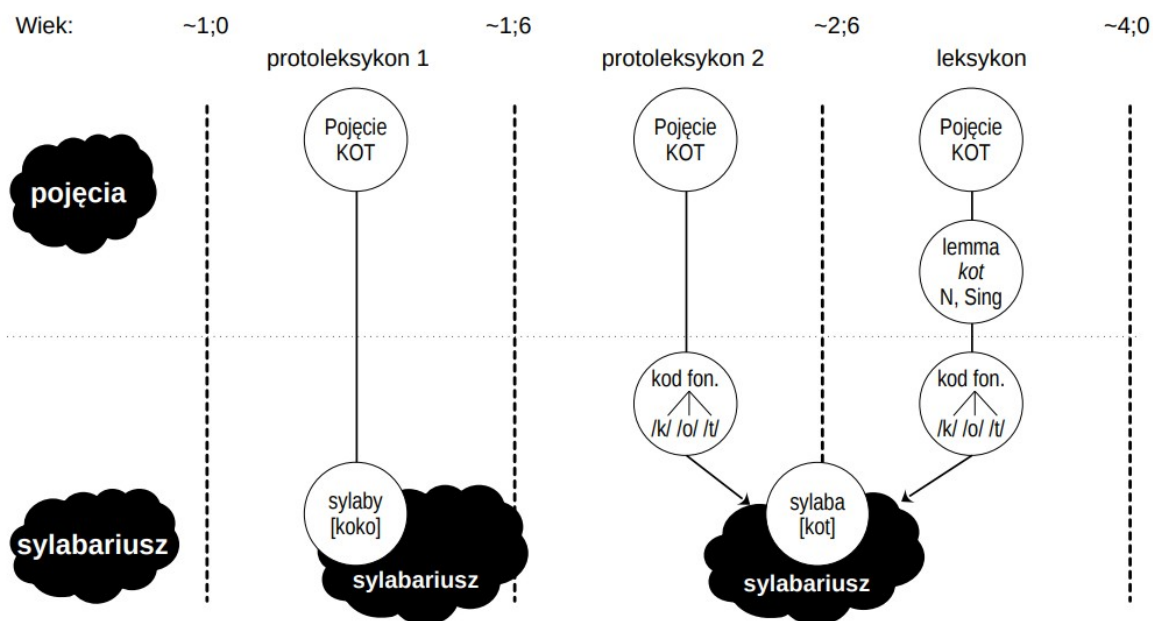
Do podobnych wniosków zdaje się dochodzić Levelt (1998), opisując proces rozwoju produkcji mowy<sup>95</sup> w początkowych latach życia dziecka. Po pierwsze, stwierdza on, że załączki obu systemów, tj. systemu retoryczno-semantyczno-syntaktycznego i fonetyczno-fonologicznego kształtują się już w pierwszych dwunastu miesiącach życia, a więc na długo zanim dziecko zacznie mówić oraz tego, że początkowo kształtują się całkowicie niezależnie od siebie. Po drugie, opisuje on proces stopniowej formalizacji połączeń między tymi systemami dzięki postępującym od 2. roku życia fonologizacji (*phonologization*) i gramatyzacji (*grammaticalization*) leksykonu umysłowego. Rozwój procesu mówienia, wyznaczony strukturalną przebudową leksykonu umysłowego przedstawia rysunek 7.

Zdaniem Levelta (1998) podstawą dla systemu retoryczno-semantyczno-syntaktycznego są zdobywane przez dzieci w pierwszym roku życia bazowe pojęcia miejsca i czasu, intencji oraz przyczynowości. Wsparcia dla tego stwierdzenia dostarczają chociażby klasyczne badania Piageta (1966) nad rozwojem inteligencji u niemowląt. Piaget dowodził, że kolejne osiągnięcia inteligencji sensoryczno-motorycznej służą strukturalizacji doświadczeń dziecka, organizując rzeczywistość w schematy stałego przedmiotu, przestrzeni, czasu i przyczynowości. Schematy te stanowią „substruktury późniejszych odpowiadających im pojęć” (Piaget i Inhelder, 1993 s. 18). Najwięcej dowodów na wczesne, pozajęzykowe nabywanie kategorii semantycznych, które stanowią fundament pojęciowy dla przyswajania słów, pochodzi z badań nad określaniem relacji przestrzennych i używaniem przez dzieci wyrażen lokatywnych (Bowerman i Choi, 2004; Slobin, 1980).

95 Levelt (1998) opiera się na modelu dostępu leksykalnego w generowaniu pojedynczych wyrazów (np. w czynności nazywania przedmiotów). Taki proces (wypowiadania pojedynczych wyrazów) modeluje WEAVER++ (Levelt i in., 1999). Levelt (1999b) rozszerza ten model, obejmując nim generowanie bardziej złożonych wypowiedzi, zgodnie z wcześniej opracowaną teorią (Levelt, 1989; por. rozdział 2). Stąd opisywany przez Levelta (1998) proces rozwojowy dotyczy przede wszystkim generowania pojedynczych wyrazów, ale można rozciągnąć na rozwój produkcji mowy „w ogóle”.

## Rysunek 7

### Zmiany rozwojowe w procesie produkcji mowy w oparciu o model Levelta



Źródło: opracowanie na podstawie: Levelt (1998)

Zbliżone stanowisko wyraża współczesne językoznawstwo kognitywne na temat archetypów pojęciowych, czyli pojęć takich jak „rzecz”, które są silnie zakorzenione w doświadczeniu fizycznym (somatycznym) i budują struktury bazowe dla bardziej złożonych pojęć (Taylor, 2007; Langacker, 2009). Ponadto – jak dodaje Levelt (1998) – jeszcze w okresie niemowlęcym dzieci przyswajają sobie również odrębne kategorie semantyczne ludzi, zwierząt, przedmiotów oraz różnego rodzaju kategorie działań. Ponownie, współczesne językoznawstwo kognitywne zgadza się z tym, że sama zdolność kategoryzacji jest pierwotnie pozajęzykowa, ale jednocześnie stanowi podłoże posługiwania się językiem, gdyż wyrazy są nazwami kategorii (Taylor, 2007). Osobną kwestią jest to, w jaki sposób te przedjęzykowe reprezentacje (*preverbal*) przekształcają się w pojęcia, których etykietowaniu służą przyswajane przez dzieci wyrazy (Mandler, 1992).

Równolegle w pierwszym roku życia dzieci doskonalą się w zarówno w odbiorze, rozpoznawaniu jak i produkcji dźwięków mowy, tworząc podwaliny pod rozwój systemu fonetyczno-fonologicznego (Levelt, 1998). Już niemowlęta w wieku około 1. miesiąca są w stanie rozróżniać fonemy /p/ i /b/, dzięki percepcji cechy „dźwięczności” głosek (Eimas i in., 1971). W kolejnych miesiącach życia ta specyfikacja słuchu w kierunku odbioru dźwięków języka typowego dla otoczenia wzrasta jeszcze bardziej. Na przykład Jusczyk i in. (1993)

wykazali, że niemowlęta w 9. miesiącu życia wykazują większą preferencję w słuchaniu słów języka ojczystego niż obcego. Poddane eksperymentowi dzieci amerykańskie dłużej skupiały uwagę na prezentowanych im nagraniach słów angielskich niż niderlandzkich; natomiast odwrotny kierunek zainteresowania obserwowano u dzieci holenderskich. Według autorów badania dokonywane przez niemowlęta rozróżnienia świadczą o rozwijaniu przez nie reprezentacji struktury fonologicznej języka ojczystego. Dzięki temu dzieci uczą się np. rozgraniczania wyrazów w ciągłym strumieniu mowy, ale też mniej więcej w tym czasie tracą uniwersalną zdolność wykrywania kontrastów fonetycznych w językach obcych (Łobacz, 2005).

Analogiczne zmiany rozwojowe zachodzą w zakresie produkcji dźwięków mowy, choć jednocześnie należy zaznaczyć wyraźne różnice czasowe w kwestii rozwoju percepcji i produkcji dźwięków mowy – percepcja wyprzedza produkcję w swoim zaawansowaniu (Łobacz, 2005; van Beinum, 2008). Według Levelta (1998) od siódmego miesiąca życia dziecka następuje intensywny rozwój sylabariusza, który wypełniają proste wzorce sylabiczne, głównie spółgłoskowo-samogłoskowe (CV), realizowane w postaci gaworzenia samonaśladowczego (powtarzalnego) takich dźwięków jak ba-ba, da-da, gi-gi. Aby jednak w tworzonych przez półroczne niemowlę dźwiękach dało się rozpoznać wzorce przypominające sylaby, musi ono nauczyć się wcześniej koordynować ze sobą cykle oddechu, ruchy krtani odpowiedzialne za fonację oraz ruchy mięśni traktu głosowego odpowiedzialne za artykulację. Według van Beinum (2008) prosta, nieprzerwana fonacja obecna jest od urodzenia i dominuje nad innymi sposobami generowania przez niemowlę dźwięków w ciągu dwóch pierwszych miesięcy życia. W tym początkowym okresie poprzez reakcje płaczu (krzyku), ssania i przełykania dziecko nabywa doświadczenia w kontroli przepływu powietrza przez nos i usta (Łobacz, 2005).

Pod koniec drugiego miesiąca życia dziecka (około 6-8. tygodnia) można zaobserwować występowanie przerywanej fonacji w ciągu jednego cyklu oddechowego (van Beinum, 2008). Nazywamy to głużeniem lub gruchaniem, gdyż dziecko wytwarza dźwięki samogłoskopodobne oraz spółgłoski interpretowane przez dorosłych jako tylnojęzykowe (welarne) (Łobacz, 2005). Od tego czasu przez kolejne tygodnie obserwuje się coraz większą zmienność intonacji, a część tworzonych przez niemowlę dźwięków powstaje przy aktywnym udziale ponadkrtaniowego traktu głosowego, stąd van Beinum (2008) uznaje, że po 10. tygodniu życia dziecka można mówić o początku rozwoju prostej artykulacji. Ograniczenia w generowaniu zróżnicowanego repertuaru dźwięków zdeterminowane są anatomią i fizjologią toru głosowego (Łobacz, 2005). Dlatego postępujące w tym zakresie

zmiany (wydłużanie toru głosowego, ząbkowanie) będą sprzyjać różnorodności wykonywanych ruchów (np. żucie), a tym samym rozszerzeniu repertuaru dźwięków generowanych przez niemowlę (van der Stelt, 2008).

Po 20. tygodniu życia dziecka jego fonacja ulega dużemu zróżnicowaniu. Okres ten nazywany jest etapem zabaw głosowych lub ekspansji, w którym także znacznie wzrastają możliwości artykulacyjne niemowlęcia (Łobacz, 2005). Przerwana fonacja łączona jest z prostymi ruchami artykulacyjnymi, tworząc formy przypominające sylaby. W kolejnych tygodniach obserwowane są postępy w artykulacji, która po 30. tygodniu znacznie się komplikuje (van Beinum, 2008). Początkowo, od około 25-30 tygodnia życia dziecka można mówić już o gaworzeniu samonaśladowczym, czyli powtarzaniu połączenia tej samej spółgłoski z samogłoską (sylaby kanoniczne CV) w ciągu typu *bababababa* (Łobacz, 2005). Zdaniem Levelta (1998) powstają wtedy pierwsze wzorce sylabiczne zasilające umysłowy sylabariusz. Początkowo różnią się one znacznie od sylab spotykanych w języku otoczenia dziecka, jednak z czasem zaczynają przypominać sylaby języka rodziców, jednocześnie coraz bardziej różniąc się od struktur artykulacyjnych typowych dla innych języków, ale nieużywanych w otoczeniu dziecka (De Boysson-Bardies i in., 1989; Levitt i Utman, 1992).

Około 40. tygodnia gaworzenie ma już charakter znacznie bardziej zróżnicowany (gaworzenie konkatenacyjne), gdyż w ciągu artykułowanych dźwięków pojawiają się różne spółgłoski i samogłoski, co tworzy konstrukcje przypominające słowa języka dorosłych (Łobacz, 2005; van Beinum, 2008). W tym czasie pojawiają się także pierwsze próby referencyjnego użycia dźwięków mowy, czyli używania tych samych dźwięków (sylaby lub kombinacje dwóch sylab) konsekwentnie w specyficznym kontekście (van der Stelt, 2008). Łobacz (2005, s. 238-239) nazywa te pierwsze znaczące, tzn. posiadające odniesienie wypowiedzi „pierwszymi prototypami wyrazów”. Uważa ona, że występują one w okresie przejściowym między gaworzeniem a używaniem wypowiedzi o charakterze językowym (pierwszy etap rozwoju fonologicznego), tj. do czasu kiedy dziecko opanuje pierwszych 50 wyrazów (około 18. miesiąca życia). Podobnie Levelt (1998) te pierwsze protowyrzy (*protowords*) uważa za pierwotne ogniwa łączące sylabariusz z fundamentalnymi pojęciami rozwiniętymi w okresie niemowlęcym. Dzięki tym połączeniom dziecko rozwija swoisty protoleksykon (*protollexicon*), obejmujący początkowo kilka, kilkanaście a z czasem kilkadziesiąt wyrażen odnoszących się do ludzi, zwierząt, przedmiotów i czynności. Podsumowanie osiągnięć w rozwoju fonacji i artykulacji przedstawia tabela 2.<sup>96</sup>

---

96 Dokładniejsze zestawienie osiągnięć rozwoju mowy w pierwszym roku życia obejmujące zmiany w „kompetencji percepcyjnej” oraz „kompetencji artykulacyjnej” zawiera praca Łobacz (2005, s. 239).

**Tabela 2***Kamienie milowe w rozwoju fonacji i artykulacji w okresie niemowlęctwa*

Stadium	Opis zmian	Wiek
1	Prosta, nieprzerwana fonacja bez ruchu artykulacyjnego w jednym cyklu oddechowym.	0 tyg.
2	Przerywana fonacja w ciągu jednego cyklu oddechowego, bez ruchu artykulacyjnego.	6 tyg.
3	Jeden pojedynczy ruch artykulacyjny w ramach jednego cyklu oddechowego, który może być połączony z (1) ciągłą lub (2) przerywaną fonacją.	10 tyg.
4	Zmniejszenie występowania ruchów artykulacyjnych i wzrost liczby dźwięków z nieprzerwaną fonacją, która zmienia się pod względem głośności, wysokości i czasu trwania.	20 tyg.
5	Wielokrotne ruchy artykulacyjne, powtarzalne lub różnorodne, podczas jednego cyklu oddechowego wraz ze zmiennością fonacyjną (gaworzenie).	26 tyg.
6	Kombinacje wszystkich typów „wypowiedzi” z poprzednich etapów oraz początek referencyjnego użycia dźwięków: formy zgodne fonetycznie i pierwsze wyrazy.	40 tyg.

*Adnotacja:* opracowano na podstawie van der Stelt (2008)

W tym miejscu warto podkreślić, że opisywany rozwój mówienia nie odbywa się w „próżni społecznej”, ale w intensywnej interakcji między dzieckiem a jego otoczeniem. Płacz, krzyk, czy śmiech służą komunikowaniu stanów emocjonalnych dziecka, na które pojawia się odpowiedź ze strony opiekunów, gdyż względnie właściwie potrafią oni zinterpretować to zachowanie jako oznakę (symptom<sup>97</sup>) danego stanu. Występuje tu wyraźna nierównowaga między kompetencjami dziecka, które jeszcze nie wie, że coś komunikuje (a przynajmniej nie wie „co” komunikuje), a zdolnościami interpretacyjnymi otoczenia, które nadaje sens

97 Milewski (1967/2004, s. 20) dzieli znaki na sześć grup ze względu na ich strukturę i funkcję. Z tego podziału korzysta Kaczmarek (1977) wprowadzając opis zmian w rozwoju zachowań komunikacyjnych i mowy dziecka. Do symptomów (oznak) zaliczane są zjawiska wywołane przez pewne przyczyny, które są znakami jedynie dla odbiorcy, który wiąże je z jakąś treścią. Są więc niecelowe i jednostronne. Sygnały z kolei są generowane przez nadawcę aby oddziaływać na odbiorcę poprzez to, że zarówno nadawca, jak i odbiorca są w stanie je odebrać i odnieść do tej samej treści. Są więc celowe dwustronne (Milewski, 1967/2004).

zachowaniom komunikacyjnym dziecka. Ta realna relacja społeczna oraz intencje dorosłych, aby odczytać (zrozumieć) zachowanie dziecka, wspiera proces rozwojowy skutkujący „przecięciem się” – używając terminologii Wygotskiego – linii mowy i myślenia. Dokładne opisanie tego zjawiska wykracza poza ramy niniejszego opracowania, jednak warto pamiętać, że opisywany tu rozwój procesu mówienia (indywidualnej realizacji mowy) ma swoje społeczne umocowanie<sup>98</sup>. Skutkiem tego, pierwotnie nieintencjonalne symptomy, z czasem nabierają charakteru intencjonalnie stosowanych sygnałów-apeli (Kaczmarek, 1977). Od około 9.-12. miesiąca życia dzieci zaczynają przejawiać wiele zachowań związanych z uwspólnianiem uwagi, które można interpretować jako rozumienie swoich i cudzych działań w kategoriach realizacji intencji. Dzieci wtedy nie tylko zaczynają działać intencjonalnie na przedmiotach, co Piaget (1966) uznaje za właściwy początek inteligencji sensoryczno-motorycznej, ale też dążą do przekazania swoich intencji innym ludziom. W tym celu w wieku około 1. roku zaczynają stosować gesty wskazujące, które według Tomasello (2002) mają charakter trójstronny, to znaczy nie tylko odnoszą się do danego przedmiotu, ale też wskazują dorosłemu na ten przedmiot, oznajmiając coś (celem jest zwrócenie uwagi) lub próbując nakłonić dorosłego do działania z tym przedmiotem. Na tle tych epizodów wspólnej uwagi i zaangażowania pojawiają się pierwsze prototypy wyrazów (protowyrazy), które nie tylko „coś” oznaczają, ale wyrażają jakąś intencję komunikacyjną. Podobnie jak w przypadku gestów mają one na celu zwrócenie uwagi na jakiś przedmiot lub zdarzenie albo wyrażają prośbę lub żądanie (mają charakter imperatywny) względem interlokutora i oznaczanego przedmiotu. Można więc o nich mówić w kategoriach aktów illokucyjnych (por. Austin, 1993; Searle, 1980). Wspólna uwaga oraz rozumienie siebie i innych jako intencjonalnych sprawców (Tomasello, 2002) są więc społeczno-poznawczą podstawą dla skrzyżowania się linii „myślenia” (rozwój systemu pojęciowego) i „mowy” (rozwój fonacji i artykulacji), które według Wygotskiego (1989/1934, s. 51) „dają początek zupełnie nowej, tak charakterystycznej dla człowieka formie zachowania się”, jaką jest mówienie.

---

98 Tomasello (2002, s. 68) zauważa, że wielu psychologów poznawczych na ogół ignoruje wpływ procesów historyczno-kulturowych oraz ontogenezy na kształt „dojrzałych form poznania ludzkiego”. Trudno nie zgodzić się z tym zarzutem również w odniesieniu do poznawczych modeli procesu mówienia (zob. rozdział 2), które opisują proces produkcji mowy po pierwsze w pewnym oderwaniu od sytuacji komunikacyjnej (paradygmaty badawcze obejmują nazywanie obrazków lub powtarzanie „łamańców językowych” zamiast badania mówienia w rozmowie), a po drugie bez uwzględnienia procesów rozwojowych odróżniających mówienie dzieci od mówienia osób dorosłych. Wyjątkiem, przynajmniej w tej drugiej kwestii, jest propozycja Levelta (1998; Levelt i in. 1999), który źródła rozdziału między systemami retoryczno-semantyczno-synaktycznym oraz fonetyczno-fonologicznym w mówieniu widzi w ontogenezie tego procesu. Pierwsza kwestia, czyli społeczne procesy kształtujące zdolność mówienia są w tym modelu pomijane. Z tego względu również w niniejszej pracy są one tylko sygnalizowane jako ważne tło dla procesu mówienia oraz rozwoju jego kontroli.

### 4.1.2 Nabywanie języka a zmiany w strukturze procesu mówienia

W okolicy 18 miesiąca życia, po opanowaniu pierwszych 50 wyrazów następuje nagły, intensywny przyrost liczby wypowiedzianych słów, określane czasem „eksplozją nazywania”. Zdaniem Levelta (1998), wraz z rozbudowywaniem protoleksykonu dziecko staje przed problemem rozróżniania coraz większej liczby podobnych sylab ze względu na to, do czego się odnoszą. Według Łobacz (2005) dziecko wchodzi wtedy w drugą fazę rozwoju fonologicznego, w której na podstawie powtarzalnych zachowań artykulacyjnych tworzy ono własne reguły fonologiczne. Pierwsza faza wiązała się z prymatem jednostek leksykalnych jako całości nad strukturą fonologiczną wyrazów. Ferguson and Farwell (1975) twierdzą, że rdzeń foniczny (*phonic core*) zapamiętanych jednostek leksykalnych i artykułowanych sylab oraz nowe dane wejściowe (nowoprzyswajane wyrazy) stają się fundamentem dla rozwoju własnej fonologii po opanowaniu 50 pierwszych wyrazów. Ich zdaniem dziecko zaczyna dokonywać generalizacji na podstawie posiadanego repertuaru wzorców artykulacyjnych i reorganizować je wedle tych zasad. Pojawiają się wtedy błędy w wypowiedziach dziecka polegające na zamianie (substytucji) głosek (np. końcówka [k] zamieniana na [t] w wyrazie *book*), asymilacji i innych procesach fonologicznych, świadczące o tym, że system produkcji mowy przeorganizowuje się z poziomu realizacji całych jednostek leksykalnych na poziom segmentowego systemu fonologicznego (Ingram, 1986; Crosbie, Holm i Dodd, 2009; por. Zarębina, 1965). Badania na dzieciach niderlandzkojęzycznych (Levelt, 1994, za: Levelt, 1998) pokazały, że początkowo ich protoleksykon składa się w większości z wyrazów, w których spółgłoski realizowane są w zbliżonych miejscach artykulacji np. spółgłoski w wyrazie [tin] są przedniojęzykowe, a w wyrazie [pom] – dwuwargowe. Z czasem jednak dziecko zaczyna zwracać uwagę na początek słowa, dowolnie zmieniając miejsce artykulacji początkowej spółgłoski, tworząc np. kombinacje [pin] lub – jeśli uwaga skierowana jest na koniec wyrazu – [pok]. W końcu dziecko wyodrębnia jądro sylaby, które można połączyć z dowolnymi spółgłoskami. Zdaniem Levelta (1998) obserwacje te są dowodem na postępującą w okresie 18-30 miesięcy fonologizację<sup>99</sup> (*phonologization*) leksykonu, czyli tworzenie przez dziecko generatywnego systemu indeksowania fonologicznego, który umożliwi rozdzielenie w sylabariuszu dużej liczby gestów sylabicznych. Zarębina (1965) zwraca uwagę, że rozwój językowy dzieci odbywa się wtedy między innymi przez zdobywanie nowych opozycji fonologicznych. Oprócz tego, że w repertuarze dziecka pojawiają się realizacje nowych fonemów (różnicowanie miejsca artykulacji spółgłosek), to komplikuje się także struktura

99 Termin „fonologizacja” jest stosowany w języku polskim na określenie procesu zmian ewolucyjnych języka w czasie, prowadzącego do tego, że dwie głoski, które wcześniej były wariantami tego samego fonemu (allofonami), stają się opozycyjne względem siebie (Polański, 2003).

syłabowa produkowanych wyrazów: poza reduplikacjami otwartych, kanonicznych sylab (CVCV) lub sylab zamkniętych (CVC), coraz częściej w mowie dziecka pojawiają się zbitki spółgłoskowe, najwcześniej jako dwuelementowe grupy spółgłoskowe w środkowej części wyrazu (CVCCV) (Łobacz, 2005). Jak zauważa Łuczyński (2019, s. 175) opanowywany przez dziecko schemat językowy początkowo jest niepodzielny wewnętrznie, holistyczny, jednak z czasem dziecko zyskuje „świadomość kompozycyjności obrazów dźwiękowych, w tym także umiejętność wyławiania z wypowiedzenia elementów składowych.”

Według Levelta (1998) zasadniczy proces fonologizacji<sup>100</sup> kończy się w pierwszym półroczu trzeciego roku życia. Nie oznacza to bynajmniej zakończenia rozwoju systemu fonologicznego. Na przykład rozwój umiejętności realizowania zbitki spółgłoskowych rozciąga się na kilka lat i będzie postępować jeszcze w okresie przedszkolnym (Łobacz, 2005). Jednak, jak zauważa Zarębina (1965, s. 90) – w trzecim roku życia system dziecka „tylko w niewielu szczegółach będzie się różnił od systemu języka otoczenia”.

O ile pierwszy rok życia dziecka nazywany bywa „okresem melodii”, drugi rok – „okresem wyrazu”, o tyle o trzecim roku życia można powiedzieć, że jest „okresem zdania”, czyli etapem rozwoju mowy, w którym dziecko zaczyna tworzyć coraz dłuższe wypowiedzi (Kaczmarek, 1977). W zasadzie już w drugiej połowie drugiego roku życia, wraz z rozwojem protoleksykonu obserwuje się występowanie wyrażen dwuwyrzowych, w których relacje między słowami mają charakter semantyczny i nie są jeszcze wyznaczone regułami składni (Kurcz, 2005), a . Zdaniem Levelta (1998) w trzecim roku życia dziecka, w strukturze jego leksykonu brakuje jeszcze reprezentacji poziomu lemma, czyli wyrazu składniowego (*syntactic word*). Wyrazy są więc łączone na zasadzie pełnionych przez nie funkcji odniesienia, gdyż w doświadczeniu dziecka często występują „(...) w sytuacjach mających wspólne cechy kontekstowe i behawioralne” (Bloom, 1980, s. 246). Występujący w tych konstrukcjach specyficzny (nielosowy) i konsekwentny szyk wyrazów zdradza, że zaczynają być one organizowane wedle relacji gramatycznych. Według Levelta (1998) mniej więcej od około wieku 2,5 lat następuje proces intensywnej gramatyzacji leksykonu<sup>101</sup>, przeorganizując go wedle formalnych kategorii gramatycznych (na kategorie leksykalne nałożone zostają kategorie gramatyczne). W ten sposób przez wzajemne przyporządkowanie systemu pojęciowego (określającego role tematyczne składników w strukturze predykatowo-

---

100 Fonologizacja prowadzi do wykształcenia się reprezentacji fonologicznej (fonologicznego kodu) mówionych wyrazów, które w procesie produkcji mowy przyjmują postać wyrazów fonologicznych (*phonological words*) (por. Levelt i in. 1999). Z perspektywy struktury znaku językowego można powiedzieć, że dzięki temu procesowi substancja dźwiękowa organizowana jest w specyficzną formę wyrażania.

101 Gramatyzację leksykonu, skutkującą rozwinięciem poziomu reprezentacji lemma można uznać za proces organizujący substancję pojęciową dzięki językowej formie planu treści.

argumentowej) i syntagmatycznego (wyznaczającego pozycje syntagmatyczne składników zdania) tworzone są kategorie części mowy (por. Frydrychowicz, 1988). Dzięki temu procesowi, używane w mówieniu wyrazy pozostają względem siebie nie tylko w relacji semantycznej, ale ich użyciem zaczynają rządzić reguły składni, co umożliwia budowanie wielowyrazowych zdań (Kurcz, 2005; Łuczyński, 2019). W tym czasie w mowie dzieci 2,5-3-letnich wzrasta liczba używanych wyrazów pomocniczych, pozwalających wyrazić relacje syntagmatyczne między poszczególnymi wyrazami treściowymi, którym odpowiadają lemmy (Levelt, 1998). Równoległe do rozwoju składni nabywanie systemu gramatycznego przejawia się w rozwoju fleksji (Łuczyński, 2019), dlatego w mowie dzieci posługujących się językami syntetycznymi obserwuje się zmianę rozwojową w kwestii używania morfemów fleksyjnych. Według Frydrychowicza (1988) końcówki wyrazów „-ki”, „-oć”, „-ka”, które w mowie osób dorosłych służą oznaczaniu wartości kategoryjnej (są formalnymi wykładnikami liczby, przypadku, czasu, rodzaju itd.), w mowie dzieci poniżej 3 lat wskazują oznaczenie wartości klasowej. Oznacza to na przykład, że początkowo wypowiedziana przez dziecko sylaba „-ki” odnosi się bezpośrednio do pojęcia „chusteczki” i nie wyraża wartości kategoryjnej tego wyrazu (liczba mnoga rzeczownik rodzaju żeńskiego), a jego wartość klasową. Wypowiedź „-ki” funkcjonuje więc pierwotnie jak morfem leksykalny, ale z czasem – w wyniku gramatyzacji leksykonu umysłowego – zaczyna pełnić funkcję morfemu gramatycznego, stając się końcówką wyrazu „chusteczki” (ale też „piłki”, „kropki” itp.) i wyrażając jego wartość kategoryjną. Można więc powiedzieć, że substancja planu treści otrzymuje właściwą sobie formę językową (gramatyczną), która organizuje połączenie planu treści i planu wyrażania. Według Slobina (1980), który analizował dane z różnych języków, zwracanie uwagi na końcówki wyrazów oraz używanie ich (oraz innych wykładników gramatycznych) w sposób umotywowany semantycznie, mają charakter uniwersalnych zasad kierujących nabywaniem gramatyki przez dzieci (zasady operacyjne A i G w zestawieniu autora).

Proces rozwojowy prowadzący do ukonstytuowania w leksykonie umysłowym poziomu lemmy przejawia się w składni i fleksji języka dziecka. Jak zauważa Łuczyński (2019) poszczególne składniki wypowiedzenia nabierają wyraźnych cech określonych części mowy z ich właściwościami fleksyjnymi. Stopniowe profilowanie formalne wyrazów w wypowiedziach powoduje, że „poszczególne składniki wypowiedzenia nabierają wyraźnych cech określonych części mowy z ich właściwościami fleksyjnymi” (Łuczyński, 2019, s. 182). To z kolei ma swoje konsekwencje dla systemu pojęciowego, którym posługuje się dziecko, gdyż pozwala na tworzenie (derywację) wyrazów z danej kategorii gramatycznej (np. rzeczownika), przy wykorzystaniu opanowanych zasad fleksji języka (użycie przyrostka

będącego wykładnikiem). Otwiera to możliwość wyrażania całkiem ogólnych pojęć, przy wykorzystaniu prostych środków morfologicznych. W badaniu Haman (2005) rzeczowniki pochodne od przymiotników i czasowników były obecne w repertuarze dzieci już w przedziale wiekowym 2;0-4;5 lat, przy czym rzeczowniki charakteryzujące (*characterizing nouns*), takie jak „biedak”, „biedaczyna”, „chudzielec” były obserwowane wcześniej i bardziej licznie niż abstrakcyjne rzeczowniki – nazwy własności (*property names*), np. „bieda”, „chudość”, „mądrość”.

Początkowo w okresie rozwoju, o którym mowa, obserwuje się przejawy różnego rodzaju trudności w konstruowaniu wypowiedzi (przerwy, powtórzenia, reperacje, falstarty, wahania). Levelt (1998) interpretuje te nie płynności, jako efekt uboczny nowych możliwości syntaktycznych, z którymi na początku dzieciom trudno jest sobie poradzić. Efektem rozwoju gramatycznego jest nie tylko wzrost długości zdań przez dodanie wyrazów funkcyjnych i zwiększenie liczby wyrazów treściowych, ale przede wszystkim podporządkowanie szyku tych wyrazów regułom składniowym. Obserwuje się wtedy zmianę w strukturze zdania i odchodzenie od tzw. szyku psychologicznego (naturalnego) odzwierciedlającego mechanizmy funkcji poznawczych (wyrażającego bezpośrednio doświadczenia sensoryczno-motoryczne) i stosowanie szyku gramatycznego, organizującego postrzeganą (i opisywaną w mowie) rzeczywistość w kategoriach językowych (Frydrychowicz, 1988). Zdaniem Tarkowskiego (1992, s. 39) to przejście stanowi „istotę strukturalizacji gramatycznej wypowiedzi, ale jednocześnie przyczynia się do nasilenia nie płynności mówienia w postaci rewizji i pauz. Według Levelta (1998) jest to nieuniknione, aby system uzyskał stabilizację, co dokonuje się w wieku około 4-5 lat. Produkcja mowy przypomina wtedy proces opisywany przez model odnoszący się do mówienia osób dorosłych (Levelt i in., 1999; Levelt, 1999b). Wielu autorów zgadza się, że mowa dzieci w tym wieku jest już na tyle ukształtowana, że można porównać ją z mową dorosłych, jednak – jak zauważa Kurcz (2005) – dziecko wchodzi wtedy w okres osiągania pełnej kompetencji językowej, który rozciąga się do połowy, a nawet końca późnego dzieciństwa (wejście w okres pokwitania). Nabywane są wtedy i utrwalane peryferyjne elementy systemu fleksyjnego (Łuczyński, 2019). Istotne zmiany rozwojowe dotyczą też świadomości metajęzykowej, która wzbogacana jest o deklaratywną wiedzę na temat struktury języka, zdobywaną w procesie formalnej edukacji szkolnej (Kwarciak, 1995). Dzięki temu dziecko uświadamia sobie związki między elementami języka, na które wcześniej nie zwracało uwagi, np. związek między imiesłowem (używanym jak przymiotnik), a czasownikiem (Łuczyński, 2019). Sprzężenie rozwoju kompetencji językowej z nauką czytania i pisania stawia przed dzieckiem także nowe problemy, np. przełożenie struktury

fonetycznej wypowiedzi na zapis graficzny (problem relacji ortofonii i ortografii). Jednocześnie dzięki postępom w rozwoju poznawczym, np. decentracji poznawczej i wzrostowi poziomu refleksyjności w działaniu (por. Piaget i Inhelder, 1993) dzieci coraz większą uwagę zwracają na kwestie stylistyczne i retoryczne swoich wypowiedzi. Zmiany te, choć istotnie kształtujące dojrzałą kompetencję językową, nie wnoszą modyfikacji do struktury procesu mówienia, która stabilizuje się pod koniec okresu przedszkolnego (por. Levelt, 1998).

## 4.2 Rozwój kontroli procesu mówienia

Zarysowany w poprzednim podrozdziale przebieg rozwoju procesu mówienia od 1. do około 5. roku życia nie byłby możliwy gdyby nie postępująca kontrola nad różnymi parametrami tego procesu, począwszy od kontroli oddechu<sup>102</sup>. Zanim jeszcze dziecko zacznie posługiwać się mową, uczy się kontroli fonacji oraz artykulacji. Jest to niezbędne, aby mówienie oparte na wibracjach fałdów głosowych i ruchach aparatu artykulacyjnego, zapewniającego podwójny rezonans było w ogóle możliwe. Wyrazem postępującej kontroli w tym względzie są kolejne stadia głużenia, zabaw głosowych i gaworzenia opisane powyżej. O ile generowane do połowy pierwszego roku życia dziecka dźwięki mogą pozostawać przede wszystkim pod kontrolą kanałów dotykowego i proprioceptywnego, o tyle gaworzenie samonaśladowcze wymaga kontroli za pomocą kanału słuchowego, kluczowego dla zdolności monitorowania produkcji mowy. Świadczą o tym przede wszystkim dane z badań nad dziećmi rozwijającymi się prawidłowo oraz posiadającymi znaczny ubytek słuchu lub głuchymi. Poza tym, jak wspomniano wcześniej, istnieją przesłanki, że gaworzenie „dryfuje” w kierunku języka rodziców, coraz bardziej upodabniając się do niego.

W kontekście kontroli procesu mówienia warto odnotować jeszcze, że pierwsze protowyrazy, poza referencją, która jest odczytywana przez otoczenie społeczne w kategoriach znaku językowego – różnią się od gaworzenia tym, że są to najczęściej dwie lub jedna sylaba, a nie ciąg wielu powtarzających się sylab (por. Łobacz, 2005). Można więc przypuszczać, że obok umiejętności uruchomienia odpowiedniego programu motorycznego, ważnym mechanizmem kontrolnym, leżącym u podłoża mówienia jest zdolność do interrupcji, czy też zahamowania artykulacji<sup>103</sup>. W rozdziale trzecim zatrzymanie mówienia

102 Badania porównawcze pokazują, że człowiek ma niespotykaną u innych naczelną zdolność świadomej kontroli oddechu, co związane jest z unerwieniem mięśni obsługujących klatkę piersiową. Pozwala to m.in. na wydłużoną fazę wydechową, która podczas mówienia może zająć nawet 90% całego cyklu oddechowego (Żywicznyński i Waciewicz, 2015).

103 Pierwsze słowa (holofrazy) nie przychodzą po fazie „bezgłosu”, czy „niemówienia”, ale następują po fazie gaworzenia, a więc umiejętność zatrzymania artykulacji w odpowiednim momencie jest ważna dla

opisane było, jako konieczne dla reperacji wykrytego w wypowiedzi błędu. Aby jednak błąd mógł zostać wykryty, konieczne jest, aby produkowane dźwięki były rzeczywiście mówieniem, tj. były znakami języka, wyrażającymi intencję komunikacyjną oraz aby nadawca rozwinął zdolność monitorowania pozwalającą na porównanie produktu końcowego mówienia (artykułowanego tekstu) z tą intencją (por. Levelt, 1989).

#### **4.2.1 Wczesne przejawy monitorowania procesu mówienia**

Jak już wspomniano, w okolicy 18. miesiąca życia dziecka, kiedy opanowało już blisko 50 wyrazów (protowyrazów) zaczyna ono nagle przyswajać coraz więcej słów, rozwijając w ten sposób swój leksykon (Levelt, 1998; Łobacz, 2005). Wtedy też odnotowuje się pierwsze przypadki, kiedy popełniony przez dziecko błąd jest przez nie poprawiany (Laakso, 2010). Clark (1978) wspomina, że pierwsze odnotowane w literaturze przedmiotu (dzienniczkach mowy) reperacje pojawiają się u dzieci na początku drugiej połowy, drugiego roku życia. Na przykład dziecko w wieku 1;8 mówi „To są bolu... malinki”, gdyż chcąc powiedzieć „to są malinki” myli się i wybiera z leksykonu słowo „borówka”, po czym jeszcze w trakcie realizacji błędnego słowa zatrzymuje artykulację i wypowiada właściwy wyraz<sup>104</sup>. Okazuje się, więc że dzieci są zdolne do korekty swoich błędów w produkcji mowy niemalże tak szybko, jak tylko zaczną mówić (Hanley i in., 2016). Chociaż ich wypowiedziom daleko jeszcze do formy wypowiedzi, jakie słyszą w swoim otoczeniu, są one zdolne do monitorowania efektów swojego mówienia i porównywania ich z intencją. W przypadku, kiedy wypowiedziane słowo nie odnosi się do tego, o czym dziecko chciało powiedzieć, jest ono w stanie dostrzec tę niezgodność i poprawić swoją wypowiedź.

Zastanawiając się na tym, skąd bierze się ta umiejętność do monitorowania, ewaluacji i korekty swoich wypowiedzi należy pamiętać, że są one środkami komunikacji, która podporządkowana jest wymianie intencji (przekazanie swojej intencji i zrozumienie intencji rozmówcy). Dzieci dążą do tego, aby rozumieć innych i być przez nich rozumianymi,

---

poprawnego i płynnego wypowiedzenia danego wyrazu, np. „mama” (a nie „mamamama”), czy „lala”. Np. Spionek (1981, s. 146, za: Tarkowski, 1992) porównuje powtarzanie sylab w jękanii klonicznym do gaworzenia: „Taka postać jękania przypomina w swoich fragmentach dziecięce gaworzenie polegające, jak wiadomo, na wielokrotnym «łańcuchowym» powtarzaniu tych samych sylab lub ich części. W swojej «czystej» postaci, pozbawionej zupełnie elementów tonicznych «jękanie kloniczne» można by uznać za częściowy nawrót do okresu gaworzenia.”

104 Podany przykład został zanotowany przez autora niniejszej rozprawy u swojej córki w wieku 1;8. W okolicach tego wieku dziecko do błędnych wypowiedzi zaczęło także czasami dodawać partykułę „nie”, sygnalizując popełnienie (i wykrzycie) pomyłki. Negacja błędu jest oznaką jego detekcji przez system monitorowania mowy i może pojawiać się w wypowiedzi bez reperacji – szczególnie u osób, które mają trudności z planowaniem wypowiedzi (np. trudności w dostępie leksykalnym u osób z afazją) (Nozari, 2020).

wzajemnie doświadczając tego dążenia także od otaczających ich ludzi. Z tego względu dzieci nie tylko poprawiają swoje błędne wypowiedzi, ale też same są poprawiane przez innych, co może rozwijać w nich pewną czujność i świadomość tego, że popełniony błąd należy poprawić, aby komunikat został właściwie odebrany (por. Clark, 1978; Tarkowski, 1992). Ma to odzwierciedlenie w autokorektach dziecka, ale także w tym, że poprawia ono wypowiedzi dorosłych, jeśli – próbując zrozumieć wypowiedź dziecka – błędnie ją parafrazują. Kaczmarek (1977, s. 58) przytacza następującą sytuację: „Gdy na przykład dziecko (1,8,3) mówi *dieczinka ńde mamusi*, a opiekun powtórzy «tak dziecinka idzie» – wówczas dziecko gorąco poprawia: *nie dietinka, dieczinka!* Zapytanie «dziewczynka?» uznaje za właściwe.” Sytuacja ta pokazuje, że dziecko aktywnie dąży do przekazania swojej intencji oraz zabiega o to, żeby ta intencja została właściwie zrozumiana. Podobne dążenie, aby rozumieć i być zrozumianym przejawia oczywiście także dorośli.

Przytoczone przykłady są więc nie tylko dowodem aktywnego monitorowania i – w razie potrzeby – korygowania przez dziecko własnych i cudzych wypowiedzi, ale wpisują się w szersze zjawisko reperacji konwersacyjnych, które są podejmowane przez interlokutorów w celu przywrócenia zaburzonego toku rozmowy. Forrester (2008), który dokumentował rozwój zdolności reperacyjnych jednego dziecka (studium przypadku własnej córki) zauważył, że bardzo wcześnie (już w pierwszych miesiącach po ukończeniu 1 r. ż.) dzieci stosują bogaty repertuar zachowań, służących poradzeniu sobie z niepowodzeniem w komunikacji (najczęściej kiedy brak jest adekwatnej odpowiedzi ze strony dorosłego na komunikat dziecka). Należą do nich zmiany dźwięku (intonacja, głośność), gestykulacja, czy łączenie obu tych elementów. Forrester (2008) zauważył również tendencję do autoreperacji konwersacyjnych (modyfikacji własnej wypowiedzi w obrębie jednej tury konwersacyjnej). To znaczy, że kiedy to możliwe, nie tylko dorośli, ale też dzieci bardziej preferują autoreperacje niż reperacje inicjowane przez współrozmówcę. Analiza rozwoju umiejętności reperacji czwórki dzieci dokonana przez Laakso (2010) pokazuje, że autoreperacje dokonywane przez dziecko w pierwszej turze konwersacji (ta sama tura, w której wystąpił błąd) pojawiają się wcześnie (po 18 miesiącu życia dziecka) i nie wymagają zwrócenia uwagi na błędną wypowiedź ze strony dorosłego interlokutora. Równoległe do nich zauważa się wiele reperacji w trzeciej turze konwersacyjnej, a więc motywowanych uwagą zwróconą przez dorosłego (najczęściej w postaci powtórzenia wypowiedzi dziecka w prawidłowy sposób z intonacją wznoszącą, sygnalizującą potrzebę upewnienia się co do zrozumienia treści wypowiedzi dziecka). Zapewne dzieje się tak dlatego, że wiele błędów (szczególnie fonologicznych) pozostaje niezauważonych przez dzieci przed trzecimi urodzinami (Jaeger,

2005). Z czasem jednak (po ukończeniu drugiego roku życia) autoreperacje wynikające z autodetekcji błędu zyskują znaczną przewagę nad reperacjami motywowanymi z zewnątrz (Laakso, 2010).

Na tle tej ogólnej umiejętności w dążeniu do osiągnięcia wzajemnego zrozumienia rozwija się zdolność monitorowania własnych wypowiedzi i poprawiania ich w wyniku wykrycia błędu lub nieścisłości<sup>105</sup>. Według Gyarmathy i Neuberger (2013) dzieci uczestnicząc w coraz większej liczbie interakcji, wraz z wiekiem zdobywają więcej doświadczenia w tym, jak sprawić, by ich wypowiedzi były dokładniejsze i bardziej zrozumiałe dla innych. Kluczową rolę w tym procesie odgrywają opiekunowie, którzy poprawiając wypowiedzi dziecka oraz korygując własne błędy mowy dostarczają dziecku zwrotne informacji do nauki i ćwiczenia samodzielnej kontroli mówienia. W ten sposób dzieci uczą się kryteriów wedle których monitorują swoją mowę, aby lepiej dostosować ją do słuchacza (Evans, 1985). Ma to odzwierciedlenie także w zachowaniu dorosłych, którzy więcej niespecyficznych próśb o doprecyzowanie wypowiedzi (pytanie „What?”) kierują do dzieci młodszych niż do starszych (Leavy, 1999). Według Manfry i in. (2016) młodsze dzieci mogą wykorzystywać prośby słuchaczy jako „katalizator” do zastanowienia się nad wypowiedzianymi słowami i korekty błędu bez konieczności pełnego zrozumienia dlaczego jest to błąd oraz bez pełnej wiedzy i umiejętności, żeby ten błąd poprawić. Salonen i Laakso (2009) uważają, że nawet jeśli dzieci bardzo wcześnie samodzielnie potrafią wykryć nieprawidłowości we własnych wypowiedziach, to nadal potrzebują informacji zwrotnej od bardziej kompetentnych rozmówców, aby rozpoznać, co należy naprawić (por. Jaeger, 2005).

Jak wspomniano dzieci wykazują umiejętność monitorowania procesu mówienia już na etapie wypowiedzi jednowyrazowych, poprawiając błędy leksykalne polegające na niewłaściwym etykietowaniu wskazywanych obiektów, zjawisk lub czynności. Z czasem – dzięki wsparciu otoczenia – wraz z rozwojem mowy i budowaniem coraz bardziej złożonych wypowiedzi, monitor staje się wrażliwy na inne aspekty realizacji języka w mówieniu. Jak zauważają Salonen i Laakso (2009) reperacje fonologiczne obok leksykalnych stanowią najwcześniejsze formy autokorekt w mówieniu małych dzieci, z kolei później, w okresie przedszkolnym obok reperacji leksykalnych największą grupę stanowią korekty składni (Laakso, 2010). Można więc powiedzieć, że w wyniku postępującej fonologizacji i

---

105 Chociaż przerwanie wypowiedzi w wyniku wykrycia błędu samo w sobie powoduje zakłócenie przebiegu rozmowy, jest dla odbiorcy mniej kłopotliwe niż otrzymanie błędnej informacji od nadawcy lub podjęcie próby doprecyzowania, co nadawca w rzeczywistości miał na myśli. Problem tego, w jaki sposób odbiorca łączy reperację nadawcy z jego wcześniejszą wypowiedzią (część wypowiedzi przed interupcją) szerzej omawia Levelt (1989). Laakso (2010) zauważa, że dla dziecka przerwanie wypowiedzi i wprowadzenie korekty może być skutecznym sposobem na zwiększenie zaangażowania interlokutora we wspólne działanie.

gramatyzacji kryteria poprawności wypowiedzi ulegają rozszerzeniu, a monitorowanie obejmuje coraz więcej aspektów formy językowej.

MacDonald i in. (2012) przeprowadzili badanie eksperymentalne, w którym osobom badanym podczas zadania wymagającego mówienia (wypowiadanie słowa „bed”) podawano zmienioną słuchową informację zwrotną. Manipulacja eksperymentalna polegała na modyfikacji formantów F1 i F2 wypowiedzianej głoski [ɛ] tak, aby brzmiała jak [æ] powodując zmianę znaczenia słyszanego słowa względem słowa wypowiedzianego („bad” zamiast „bed”). W badaniu wzięły udział osoby w 3 grupach wiekowych: 2- i 4-latki oraz osoby dorosłe (studenci w wieku około 19 lat). Podczas gdy osoby dorosłe i dzieci w wieku przedszkolnym podejmowały działania kompensacyjne, polegająca na zmianie generowanej samogłoski w kierunku przeciwnym do perturbacji<sup>106</sup>, dwulatki nie dokonywały takiej modyfikacji w swoich wypowiedziach<sup>107</sup>. Autorzy badania podają dwa możliwe wyjaśnienia tej różnicy w reakcjach. Po pierwsze, możliwe jest, że dzieci nie polegają tak bardzo na informacji zwrotnej płynącej z percepcji własnego wykonania, a raczej korzystają z zewnętrznych źródeł monitorowania w postaci reakcji osób, z którymi rozmawiają. Według autorów badania (MacDonald i in., 2012) opiekunowie często powtarzają i odzwierciedlają wypowiedzi dziecka, co może wspierać z zewnątrz słuchową ocenę poprawności jego własnej wypowiedzi (por. Laakso, 2010). Druga interpretacja podkreśla zmiany rozwojowe w obrębie kształtu toru głosowego postępujące w pierwszych latach życia, które powodują dużą niestabilność w artykulacji, co z kolei może przekładać się na mniejszą wrażliwość monitora słuchowego na zmienność akustyczną generowanych głosek. Chociaż dzieci wczesnie tworzą umysłowe wzorce głosek, to jednocześnie gwałtowne powiększanie leksykonu na przestrzeni 2. i 3. roku życia utrudnia wytworzenie dokładnych wzorców artykulacyjnych (Łobacz, 2005). Ponadto, jak dodają badacze (MacDonald i in., 2012), mimo iż dzieci bardzo wczesnie dokonują precyzyjnych rozróżnień dźwięków mowy (por. Jusczyk i in., 1993), kontrola słuchowa własnej mowy może kierować się innymi kryteriami (ma inne wymagania) niż percepcja mowy innych ludzi<sup>108</sup>. Na przykład w badaniu Jaeger (2005) dzieci w wieku dwóch lat poprawiały mniej niż 40% błędów fonologicznych i ponad 60% błędów leksykalnych, a w wieku 4 lat – odpowiednio ponad 60% błędów fonologicznych i ponad 75% błędów leksykalnych.

---

106 Perturbacja polegała na podwyższeniu F1 o 200 Hz i obniżeniu F2 o 250 Hz, więc działania kompensacyjne polegały na takiej modyfikacji toru głosowego aby zawyżony F1 uległ obniżeniu, a zaniżony F2 został podwyższony, dając w efekcie poprawną głoskę [ɛ] (MacDonald i in., 2012).

107 Warto tu odnotować, że badanie ukończyło tylko około połowy pierwotnie zrekrutowanych dzieci w tym wieku, co świadczy, że procedura była bardzo wymagająca dla młodszych uczestników.

108 Podobne zastrzeżenie czyni Roelofs (2020a) w odniesieniu do porównania działania systemu odbioru mowy i korzystającego z tego systemu monitora produkcji mowy (zob. rozdział 3).

Z drugiej strony warto zauważyć, że dzieci 4-letnie w cytowanym badaniu MacDonald i in., (2012) prezentowały kompensację artykulacyjną zbliżoną do reakcji dorosłych (por. Lind i in., 2014), co oznacza, że skutecznie wykorzystywały percepcję własnej mowy w ocenie realizacji głosek. Zmianę tę można przypisać postępującej fonologizacji leksykonu umysłowego w 2. i 3. roku życia dziecka (Levelt, 1998), która modyfikuje proces produkcji mowy (rozwój kodowania fonologicznego), przez co wymusza zwiększenie wrażliwości monitora na cechy artykułowanych dźwięków, które wcześniej nie podlegały monitorowaniu. Co ciekawe Salonen i Laakso (2009), w przeciwieństwie do licznych błędów fonologicznych i ich reperacji odnotowanych u dzieci 4-letnich przez Jaeger (2005), raportują, że reperacje fonologiczne u obserwowanych przez nich dzieci pojawiały się stosunkowo rzadko. Interpretują oni ten fakt w ten sposób, że mowa czteroletnich dzieci jest już dobrze rozwinięta pod względem fonologicznym i mniej jest błędnych realizacji dźwięków mowy<sup>109</sup>.

#### **4.2.2 Rozwojowa niepełność mowy a kontrola procesu mówienia**

W trzecim roku życia wraz z budowaniem bardziej złożonych zdań wzrasta także wrażliwość funkcji monitorowania na formę gramatyczną tworzonych wypowiedzi (Laakso, 2010; por. Forrester, 2008). Na przykład Snyder (1914, za: Clark, 1978, s. 24) odnotował występowanie u dziecka 2,5 letniego reperacji obejmujących:

- (1) dodanie do wypowiedzi przydawki (*modifier*) „Dat water--dat dirty water”;
- (2) zmianę słowa „Might take paddle out boat--might take paddle out canoe”;
- (3) zmianę szyku wyrazów „Down sand beach I been--I been down sand beach<sup>110</sup>.”

Dla dzieci 4-letnich w badaniu Salonen i Laakso (2009) obok korekt leksykalnych, rewizje składni stanowiły już największą grupę reperacji.

Dzięki postępującej gramatykalizacji forma wypowiedzi dziecka zostaje wzbogacona o kategorie gramatyczne, powodując powiązanie elementarnych klas pojęciowych (wykształcających się jeszcze przed rozpoczęciem mówienia na bazie sensoryczno-motorycznych doświadczeń) z podstawowymi pozycjami syntagmatycznymi (Frydrychowicz, 1988). Dzięki temu kategoryzacja świata obejmuje nie tylko podział na osoby, przedmioty, zjawiska, czy czynności, ale też na części mowy, pozwalające na wtórne (tj. dokonywane za

---

109 Warto odnotować fakt, że w obu badaniach, które miały charakter longitudinalny analizowano wypowiedzi tylko kilkorga dzieci, co utrudnia generalizację wyników. Ponadto inne cele przyświecały obu projektom. Podczas gdy Jaeger (2005) badała dziecięce przejęzyczenia (i dodatkowo ich autokorekcie), Salonen i Laakso (2009) koncentrowały się na reperacjach dokonywanych przez dzieci samodzielnie lub z pomocą dorosłego interlokutora (por. Laakso, 2010).

110 Podane przykłady można przetłumaczyć następująco: (1) „Ta woda -- ta brudna woda”; (2) „Może popłynąć łodzią -- może popłynąć kajakiem”; (3) „Na piaszczystej plaży byłem -- byłem na piaszczystej plaży (Snyder, 1914, za: Clark, 1978, s. 24)

pomocą formalnych pozycji składniowych) zrelacjonowanie względem siebie podstawowych pojęć (np. *ktoś robi coś za pomocą czegoś*). Proces ten zmienia strukturę leksykonu umysłowego (Levelt, 1998). Po pierwsze, konstryuuje się poziom reprezentacji lemma, zawierający informacje semantyczne (odnoszące się do pojęć leksykalnych) i składniowe (odnoszące się do kategorii gramatycznych). Po drugie, również poziom konceptualny ulega przekształceniom, gdyż wykształcające się pojęcia leksykalne (*lexical concepts*) (którym odpowiadają wypowiedziane wyrazy) wchodząc ze sobą w relacje tematyczne (przyjmując role semantyczne w tematycznej strukturze zdania) zawierają informacje o sposobie ich wyrażenia w kodowaniu gramatycznym (Levelt, 1989; por. Leavy, 1999). Proces tych wewnętrznych przekształceń ma odzwierciedlenie w procesie mówienia. Dzieci zaczynają budować nie tylko bardziej złożone zdania, ale używają w tym celu więcej wyrazów pomocniczych (funkcyjnych) oraz morfemów fleksyjnych (wyrażających wartość kategoriałną wyrazów), a także dostosowują szyk zdania do reguł składniowych (Frydrychowicz, 1988). Pociąga to za sobą jednocześnie rozszerzenie kryteriów poprawności pod względem, których tworzona jest wypowiedź. Dwustopniowy proces selekcji leksykalnej oraz konieczność objęcia monitorowaniem dłuższych i bardziej złożonych wypowiedzi ma wpływ na działanie mechanizmów kontroli mówienia. Jak zauważa Levelt (1998) dzieci początkowo nie są w stanie zapanować nad poprawnym wykorzystaniem nowych umiejętności składniowych (świadomość istnienia reguł składniowych przy trudności w ich stosowaniu), dlatego w tym czasie pojawia się nasilenie nie płynności w postaci interupcji, reperacji, powtórzeń i wahań. Z czasem jednak, dzięki nabywaniu sprawności w wykorzystaniu tych nowych środków gramatycznych w mówieniu zakłócenia płynności mowy zmniejszają się. Sabin i in. (1979) analizując różne rodzaje wypowiedzi osób w wieku od okresu przedszkolnego do dorosłości stwierdzili, że wzrastająca gramatyczność spontanicznej mowy w kolejnych stadiach rozwojowych odbija się w zmniejszeniu liczby zakłóceń płynności (falstarty, pauzy wypełnione lub powtórzenia) oraz we wzroście liczby dodatkowych uwag podawanych „w nawiasie”. Co istotne w tym kontekście, zakłócenia płynności mówienia u dzieci przedszkolnych dotyczą bardziej wyrazów funkcyjnych niż treściowych<sup>111</sup>, podczas gdy u dorosłych osób jękających się, nie płynność (główny objaw jękania) dotyczy przede wszystkim realizacji słów niosących znaczenie (Hartsuiker, Kolk i Lickley, 2005; por. Evans,

111 Warto w tym miejscu zwrócić uwagę, że niektóre nie płynne realizacje słów funkcyjnych można przypisać trudnościom w kodowaniu gramatycznym (por. Hartsuiker i Notebaert, 2010), ale część z nich to efekt trudności w kodowaniu fonologicznym. Jak wynika z analizy Wijnena (1992) dzieci w wieku 2-3 lat popełniają jeszcze wiele błędów w strukturze fonologicznej wyrazów pomocniczych, co jest rzadko spotykane u dorosłych. Z tego względu niektóre nie płynności (np. powtórzenia głosek lub sylab) mogą być efektem ukrytych reperacji tych błędów (por. Postma i Kolk, 1993). Jednak w badaniu Salonen i Laakso (2009) dzieci 4-letnie wykazywały na tyle wysoki poziom rozwoju fonologicznego, że stosunkowo rzadko musiały poprawiać fonologię swojej mowy.

1985; Tarkowski, 1992). Ponadto, jak wykazali Hollister i in. (2017), w przypadku dzieci przedszkolnych, u których objawy jąkania po czasie ustąpiły, miało to większy związek z wymiernym postępowaniem w rozwoju gramatycznym niż z samym wiekiem metrykalnym<sup>112</sup>.

Odnosząc zaburzenia płynności mowy do mechanizmów kontroli mówienia można je potraktować, jako przejaw niedojrzałości mechanizmów programowania mowy, które w wyniku zwiększenia obciążenia nie są w stanie poradzić sobie z tworzeniem płynnej wypowiedzi według nowych wytycznych. Według Tarkowskiego (1992) prawie u każdego dziecka w okresie, kiedy intensywnie rozwija ono język występują różne objawy niepłynności mówienia, które są rezultatem braku synchronizacji między różnymi poziomami programowania mowy. Autor ten uważa, że chwilowe trudności w swobodnym przechodzeniu od jednej do drugiej informacji (niepłynność semantyczna) wiąże się ze zwiększonym występowaniem pauz wypełnionych. Z kolei powtarzanie wyrazów funkcyjnych (głównie przyimków i spójników) jest według niego objawem trudności w swobodnym przechodzeniu z jednej struktury gramatycznej do następnej (niepłynność syntaktyczna). Odmienne źródło różnych rodzajów niepłynności potwierdza badanie Hartsuikera i Notebaert (2010), w którym manipulowali oni łatwością dostępu leksykalnego do nazw opisywanych obrazków<sup>113</sup>. W warunku eksperymentalnym, w którym selekcja leksykalna była utrudniona z powodu występowania różnych określeń na dany przedmiot, częstość pauz wypełnionych oraz autokorekt była większa niż w warunku łatwego dostępu leksykalnego, wynikającego z jednej oczywistej nazwy prezentowanego przedmiotu. Co więcej, zwiększona trudność w selekcji leksykalnej na tym poziomie nie generowała zwiększonej ilości powtórzeń (dźwięków, sylab, wyrazów, czy części fraz). Z kolei wypowiedzi wymagające użycia rzadszego określnika „het” w porównaniu z wypowiedziami zawierającymi powszechny określnik „de” prowokowały większą liczbę autokorekt i powtórzeń, ale nie pauz wypełnionych. Innymi słowy, badacze dowiedli, że trudności w selekcji wyrazu pomocniczego (określnik o niskiej częstości użycia) prowadzą do innego wzorca niepłynności niż trudności w dostępie leksykalnym wyrazów treściowych (Hartsuiker i Notebaert, 2010).

Okazuje się więc, że trudności w płynnym przechodzeniu kolejnych etapów produkcji mowy, od konstruowania przekazu przedwerbalnego do przygotowywania programu

---

112 Rozwój gramatyczny nie jest jednak jedynym czynnikiem odpowiedzialnym za niepłynność mówienia w okresie przedszkolnym, gdyż dzieci jąkające się oraz te, u których objawy ustąpiły nie różniły się zasadniczo pod względem rozwoju gramatycznego i jego poziom nie miał związku z niepłynnością mówienia u dzieci jąkających się (Hollister i in., 2017).

113 Obrazki, których położenie względem siebie opisywali uczestnicy badania, przedstawiały obiekty o różnym poziomie zgodności nazwy (obiekty o jednej nazwie vs obiekty, które można określić na kilka sposobów). Ponadto manipulowano w badaniu rodzajem wypowiedzianych rzeczowników, co wiązało się z koniecznością wypowiedzenia powszechnego (w języku niderlandzkim) określnika „de” lub znacznie rzadziej stosowanego określnika rodzaju nijakiego „het” (Hartsuiker i Notebaert, 2010).

motorycznego dla artykulatora mogą skutkować różnego rodzaju nie płynnościami. Spotyka się je w mowie spontanicznej osób nie przejawiających żadnych zaburzeń językowych i są przejawem naturalnej dynamiki procesu mówienia, jednak ich nasilenie najczęściej stanowi objaw zaburzeń takich jak jąkanie albo gielkot. U dzieci duże nasilenie nie płynności mówienia przypada na okres najintensywniejszego rozwoju języka i z czasem – u większości osób - przemija (Tarkowski, 1992).

Jak zauważają Sabin i in. (1979) dzieci, w przeciwieństwie do osób dorosłych, mają trudność z jednoczesnym mówieniem i myśleniem, dlatego mówienie i planowanie kolejnej wypowiedzi nie odbywa się u nich w pełni równolegle, ale raczej seryjnie, co rodzi zakłócenia w płynnym generowaniu tekstu. Z czasem jednak wraz z postępującą organizacją procesu produkcji mowy zakłócenia te zostają zredukowane. To zwiększenie płynności mowy można uznać – jak wcześniej zauważono – za świadectwo rozwoju gramatycznego (i fonologicznego)<sup>114</sup>. Jednocześnie jest ono dowodem dojrzewania mechanizmów kontroli mowy, które koordynują działaniem poszczególnych poziomów produkcji mowy i zabezpieczają przed wystąpieniem błędów lub usuwają ich skutki. Na rozwojową nie płynność mówienia można bowiem spojrzeć nie tylko jako na przejaw niedojrzałości systemu produkcji mowy (np. Sabin i in., 1979), ale też jako dowód działania mechanizmów kontrolnych w mówieniu. Przykładem tego są reperacje<sup>115</sup>, które dokonywane są w sytuacji wykrycia w wypowiedzi błędu lub innej nieprawidłowości (zob. rozdział 3.), co wiąże się z zatrzymaniem artykulacji (zakłócenie toku mówienia) oraz jej wznowieniem, realizującym korektę popełnionego błędu (przywrócenie płynności w toku mówienia). Do reperacji zaliczane są przede wszystkim korekty polegające na zastąpieniu błędnej realizacji poprawnym elementem. Jest to najmniej kontrowersyjny rodzaj reperacji, gdyż wyartykułowana zostaje nie tylko korekta, ale też błędna (pod względem poprawności językowej i/lub zgodności z intencją nadawcy) wypowiedź. Jak wspomniano w rozdziale 3., wiele błędów może jednak zostać wykrytych przed artykulacją, dzięki monitorowaniu wewnętrznemu, co skutkuje wystąpieniem interupcji mowy zanim artykulacja błędnego wyrazu zostanie dokończona lub w ogóle zanim się rozpocznie (Levelt, 1983; 1989; Postma i Kolk, 1993; Hartsuiker i Kolk, 2001). Z tego względu do przejawów monitorowania zaliczane są falstarty i powtórzenia. Pierwszą grupę zakłóceń stanowią wypowiedzi (wyrazy), które zostały przerwane w trakcie artykulacji. Czasami badacze biorą pod uwagę jedynie poprawne wypowiedzi, które zostały

114 Powtórzenia mniejszych elementów (głosek, sylab) w większym stopniu sygnalizują trudności w tworzeniu planu fonetycznego wypowiedzi, natomiast powtórzenia większych członów (wyrazów, części fraz itp.) bardziej odzwierciedlają trudności w kodowaniu gramatycznym (Postma i Kolk, 1993; Wijnen, 1992).

115 Wspominani Hartsuiker i Notebaert (2010) wykazali, że reperacje są tym rodzajem nie płynności, który pojawia się niezależnie od tego, czy trudność w produkcji mowy dotyczy selekcji wyrazu treściowego, czy funkcyjnego.

zatrzymane przed ich pełnym wyartykułowaniem (np. „ka”, kiedy słowem zamierzonym było „kapitan”), traktując je, jako przejaw wewnętrznego monitorowania (Roelofs, 2005a). Niekiedy jednak badacze klasyfikują jako falstarty nagłosowe fragmenty wyłącznie błędnych wypowiedzi: na przykład Gyarmathy i Neuberger (2013) analizowali niepełności mowy dzieląc je na *false starts* (niedokończone substytucje leksykalne, np. „ko-” zamiast „owca”) i *false words* (pełne substytucje leksykalne, np. „koza” zamiast „owca”).

Jeśli niedokończony początek właściwego słowa zostanie ponownie wypowiedziany (raz lub kilka razy) w całości, można mówić o powtórzeniu (np. „ko-koza”, „pa-pa-pająk”). Powtórzenia (*repetitions*) głosek, sylab, morfemów, całych wyrazów lub (części) fraz zaliczane są w literaturze przedmiotu do reperacji (Levelt, 1983; Evans, 1985; Postma i Kolk, 1993; Hartsuiker i in., 2005; Manfra i in. 2016), a czasem traktowane jako odrębna kategoria niepełności (Hartsuiker i Notebaert, 2010). Według Manfry i in. (2016) reperacje takie jak powtórzenie wyartykułowanego słowa podczas poszukiwania kolejnego właściwego słowa są oznaką samokontroli mówienia, a ich badanie daje możliwość oceny, w jaki sposób i dlaczego dzieci monitorują własną produkcję mowy.

### **4.2.3 Przejawy rozwoju kontroli procesu mówienia w okresie dzieciństwa i adolescencji**

Jak już wspomniano, pierwsze reperacje są rejestrowane w wypowiedziach dzieci w drugiej połowie drugiego roku życia (Clark, 1978; Laakso, 2010). Jaeger (2005) analizując przejęzyczenia czwórki dzieci w wieku 1-5 lat badała też jak często poprawiają one tego rodzaju błędy. Według zgromadzonych przez nią danych dzieci rzadziej korygują swoje błędy niż dorośli, ale odsetek poprawianych przez nie przejęzyczeń rośnie wraz z wiekiem: przed drugimi urodzinami dzieci poprawiały 49% swoich błędów; podobnie w wieku 2 lat – 45%, następnie liczba ta wzrosła do 55% w wieku 3 lat oraz 67% w wieku 4 lat i nieznacznie spadła do 63,6% w wieku 5 lat. Zdaniem Jaeger (2005) monitorowanie zewnętrzne (w badaniu analizowane były korekty pełnych błędów – ang. *overt repairs*) u dzieci 5-letnich nie osiąga jeszcze poziomu sprawności obserwowanego u dorosłych (73% poprawionych błędów), choć zbliża się do niego.

Zgodnie z tym rozwojowym trendem, w badaniu Evans (1985) dzieci w wieku szkolnym (2. klasa) dokonywały więcej reperacji swoich błędnych wypowiedzi w porównaniu z dziećmi w wieku przedszkolnym. W obu grupach wiekowych największy odsetek reperacji stanowiły powtórzenia. Podobnie w badaniu Manfry i in. (2016) powtórzenia były

najliczniejszą grupą nie płynności zaklasyfikowanych jako reperacje w wypowiedziach dzieci w wieku przedszkolnym. Według Evans (1985) wzrost liczby reperacji nie jest efektem większej liczby słów i dłuższych wypowiedzi (a tym samym większej liczby błędów) u starszych dzieci, ale efektem wzrastających umiejętności monitorowania mowy. Zmiana ta, ma charakter nie tylko ilościowy, ale też jakościowy. Evans (1985) donosi, że wraz z wiekiem zwiększa się liczba tzw. „odroczeń” (*postponements*), czyli autokorekt wypowiedzi niezawierających ewidentnego błędu, ale wymagających poprawienia dla uzyskania większej spójności i kompletności prezentowanych informacji (choć w analizowanym przez nią zbiorze reperacje tego typu stanowiły najmniejszą grupę autokorekt). Można odnieść je do kategorii reperacji dostosowujących (*appropriateness repairs*), które – według Levelta (1989) – są efektem bezpośredniego dostępu monitora do efektów pracy konceptualizatora i weryfikacji przekazu przedwerbalnego pod względem jego odpowiedniości wobec intencji, wiedzy nadawcy oraz kontekstu. Według Evans (1985) są one najbardziej skomplikowanym rodzajem reperacji, gdyż dostosowują daną wypowiedź do szerszego kontekstu (np. wymagają odniesienia do tematu całej konwersacji), angażując monitorowanie retrospektywne oraz prospektywne. Obecność tego typu reperacji już u dzieci 4-letnich potwierdzili również Salonen i Laakso (2009). Wśród najpowszechniejszych w tej grupie wiekowej rewizji leksykalnych i syntaktycznych, część nie obejmowała wypowiedzi błędnych, a jedynie takie które poprawiane były ze względu na możliwość bardziej czytelnej (z punktu widzenia odbiorcy) przekazania treści. Według autorów reperacje tego typu nie są motywowane językowo, a społecznie, gdyż uwzględniają perspektywę odbiorcy. Salonen i Laakso (2009) odnotowali także, że dzieci w wieku przedszkolnym – podobnie jak dorośli – dokonują reperacji po uprzednim zatrzymaniu artykulacji, ale jeszcze w tej samej turze konwersacyjnej, w której pojawił się błąd (korekty są więc efektem działania własnych mechanizmów monitorowania i kontroli mowy nadawcy, bez konieczności wskazania błędu przez odbiorcę) oraz często używają zwrotów i partykuł sygnalizujących konieczność naprawy<sup>116</sup> (por. Laakso, 2010).

Oprócz ewidentnie interakcyjnego charakteru reperacji, służą one przede wszystkim realizacji intencji nadawcy (Bręński, 2020). To znaczy, osoba mówiąca dokonuje reperacji nawet jeśli rozmówca nie sygnalizuje trudności ze zrozumieniem komunikatu. U młodszych dzieci autoreperacje obserwowane są w rozmowach z innymi ludźmi (gdyż mowa występuje wyłącznie w tym kontekście), ale w wieku przedszkolnym dzieci dokonują korekt swoich

---

116 W każdym języku występują różne zwroty ukierunkowujące uwagę rozmówcy na popełniony błąd oraz konieczność wprowadzenia korekty, np. w języku polskim stosowana jest partykuła „nie” lub zwrot „nie to miałem na myśli” albo formuła „to znaczy” itp. (por. Levelt, 1989; Salonen i Laakso, 2009; Nozari, 2020).

wypowiedzi zarówno w mowie kierowanej do innych, jak i w mowie prywatnej, np. podczas rozwiązywania zadania problemowego (Manfra i in., 2016). Oznacza, to że monitorowanie mowy i korygowanie błędów nie występuje wyłącznie w kontekście społecznej komunikacji, ale staje się funkcją kontrolną procesu mówienia w każdych okolicznościach, w których proces ten jest aktywnie realizowany. Jedno z nielicznych badań funkcji monitorowania błędów leksykalnych (semantycznych) i ich reperacji w warunkach eksperymentalnych u dzieci przeprowadzili Hanley i in. (2016). Ich zdaniem literatura bazująca na analizie spontanicznych przejęzyczeń dostarcza cennych informacji o rozwoju monitorowania mowy przez dzieci, ale może przedstawiać niedoszacowany obraz leksykalnych błędów semantycznych (często nie łamiących zasad języka) względem błędów fonologicznych, czy syntaktycznych. Badacze chcieli więc sprawdzić jak dzieci w wieku 5-7 lat monitorują i kontrolują swoją mowę poprawiając pojawiające się przejęzyczenia w zadaniu opisywania obrazków. Zgodnie z przewidywaniami wraz z wiekiem spadała liczba popełnianych przez dzieci błędów semantycznych i jednocześnie rósł odsetek dokonywanych korekt tych błędów. Istotnym predyktorem odnotowanych postępów była wzrastająca wraz z wiekiem dojrzałość organizacji leksykonu umysłowego<sup>117</sup> i umiejętności wydobywania z niego odpowiednich jednostek leksykalnych, którą badacze ocenili na podstawie poprawności odpowiedzi w niezależnym zadaniu nazywania obrazków.

Bezpośrednio sprawność monitorowania w kontekście innym niż dokonywanie korekt wypowiedzi badana była przez Sasisekaran i Weber-Fox (2012). Badaczki sprawdzały jak sprawnie dzieci i młodzież w wieku 7-13 lat potrafią monitorować wypowiedzi pod względem realizacji fonemów oraz rytmu (por. Wheeldon i Levelt, 1995). Ich badanie wykazało, że monitorowanie wewnętrzne wypowiedzi (bezgłośnie nazywanie obrazków) w poszukiwaniu określonych elementów subleksykalnych zwiększa swoją sprawność w toku rozwoju dziecka i że monitorowanie pojedynczych fonemów jest dla dzieci trudniejsze niż monitorowanie rymu (monitorowanie fonemów zajmuje więcej czasu).

#### **4.2.4 Kontrola procesu mówienia w perspektywie rozwojowej**

Reasumując dotychczasowe ustalenia, dzieci przejawiają zdolność monitorowania procesu mówienia, od kiedy tylko wkroczą w fazę komunikacji werbalnej. Już wcześniej zdradzają zdolności do monitorowania przebiegu i efektów konwersacji z dorosłymi,

---

<sup>117</sup> W badaniu nie pada termin „leksykon” (*lexicon*) tylko „lexical retrieval system” oraz „semantic-lexical part of the production system”, co wynika z oparcia badania na koneksjonistycznym modelu mówienia, który system produkcji mowy przedstawia w postaci sieci węzłów (por. Dell, 1986; Nozari i in., 2011).

dokonyjąc reperacji swoich komunikatów. Dorośli (i ogólnie otoczenie społeczne, np. starsze rodzeństwo) odgrywają ważną rolę w tym procesie jako odbiorcy komunikatów dziecka, ale też źródło informacji zwrotnej na temat efektów mówienia. Bezpośrednio poprawiając dziecko albo powtarzając jego własną (błędną) wypowiedź, ale w sposób poprawny, nie tylko przekazują informację o regułach języka, ale też modelują to, w jaki sposób dokonywane są reperacje oraz informują, jakie kryteria wypowiedzi powinny podlegać monitorowaniu. Niemniej jednak dzieci bardzo szybko prezentują zdolność dokonywania autoreperacji, co więcej preferują ten rodzaj korekt i już w 3. roku życia dominują one nad korektami dokonywanymi przy pomocy innych osób (Laakso, 2010).

W okresie wczesnego i średniego dzieciństwa zachodzą istotne zmiany w nabywaniu kompetencji językowej przez dzieci, przyczyniające się do zmian rozwojowych związanych z kontrolą mówienia. Z jednej strony rozwijanie systemu fonologicznego i gramatycznego idzie w parze z osiąganiem większej sprawności w monitorowaniu mowy, ale z drugiej stanowi wyzwanie dla monitora, gdyż wypowiedzi dziecka komplikują się, wydłużają, co zwiększa liczbę kryteriów oceny poprawności wypowiedzi. Wraz z bardziej rozbudowanymi i skomplikowanymi wypowiedziami zwiększa się okazja do popełnienia błędu, ale też zwiększa się sprawność dziecka pod względem wykrywania oraz reperacji tych błędów (Rispoli, 2003). Początkowo odbija się to znacznie na płynności mowy w postaci zaburzeń toku mówienia pod wieloma względami przypominających jąkanie (Tarkowski, 1992). Mówienie przerywane jest pauzami, zawahaniem, powtórzeniami i rewizjami dotyczącymi różnych elementów strukturalnych języka (od pojedynczych głosek do fraz). Część tych niepłynności jest efektem trudności w formułowaniu przekazu przedwerbalnego lub w dostępie leksykalnym, kodowaniu gramatycznym i morfofonologicznym, co świadczy o słabości procesów kontroli nad programowaniem mowy. Z drugiej strony część niepłynności, a szczególnie rewizje (Rispoli, 2003) których częstość wzrasta między 2-5 rokiem życia, świadczy o aktywnym monitorowaniu procesu mówienia przez dzieci w wieku przedszkolnym. Zdolność monitorowania, a w tym wykrywania błędów oraz ich skutecznej reperacji wzrasta na przestrzeni średniego i późnego dzieciństwa, co początkowo przynosi wzrost interupcji (Evans, 1985) oraz zjawisk takich jak falstarty (Gyarmathy i Neuberger (2013), ale jednocześnie spada częstość i długość pauz niewypełnionych (Sabin i in., 1979), które bardziej związane są z trudnościami w programowaniu niż służą monitorowaniu mowy (Rispoli, 2003). Można więc ogólnie powiedzieć, że w dzieciństwie zmiana ulega charakter niepłynności mowy, gdyż wraz z rozwojem językowym zmniejsza się częstość niepłynności wywołanych trudnościami w programowaniu mowy, ale wraz ze zwiększającą się kontrolą

nad produkowanym tekstem zwiększa się liczba interupcji i reperacji. Dotyczą one nie tylko błędnych wypowiedzi, ale też korekt dostosowujących, służących bardziej precyzyjnemu i zrozumiałemu dla słuchacza wyrażeniu intencji (Evans, 1985; Levelt, 1989). Zdaniem Evans (1985) związek częstości reperacji i wieku ma charakter odwróconej funkcji U-kształtnej, gdyż wraz ze wzrostem umiejętności monitorowania wzrasta też częstość autonapraw. Następnie jednak wraz z rozwojem umiejętności językowych związanych z planowaniem mówienia i organizacją myśli korekty są mniej potrzebne, a ich występowanie staje się rzadsze. Sprzyja to zwiększeniu płynności mowy w okresie szkolnym, podobnie jak spadek częstości pauz niewypełnionych<sup>118</sup>, co przekłada się na stopniowy wzrost tempa mówienia i jego najwyższą wartość w okresie adolescencji (Sabin i in., 1979).

W badaniach eksperymentalnych u dzieci w wieku szkolnym obserwuje się tendencję rozwojową polegającą na coraz skuteczniejszym monitorowaniu mowy. Hanley i in. (2016) wykazali, że wraz z wiekiem (a przede wszystkim rozwojem mechanizmów selekcji leksykalnej) spada tendencja do popełniania błędów semantycznych, natomiast wzrasta odsetek reperacji dokonywanych na tych błędach. Z kolei badanie Sasisekaran i Weber-Fox (2012) wykazało wzrastającą zdolność do monitorowania wewnętrznego wypowiedzi (nazw obrazków) pod względem ich struktury subleksykalnej, poza kontekstem detekcji błędów.

Ostatecznie po przejściu okresu adolescencji, u osób dorosłych obserwuje się rzadsze niż w dzieciństwie popełnianie błędów (Wijnen, 1992) oraz osiągnięcie wysokiej skuteczności w ich wykrywaniu i reperacji (por. Jaeger, 2005). Jest to spójne z obserwacją Sabina i in. (1979), że w porównaniu z młodszymi grupami wiekowymi, u dorosłych spada liczba niepełności w postaci powtórzeń, pauz wypełnionych oraz falstartów. Podobnie spadek częstości falstartów oraz skrócenie się fazy edycji między falstartem a reperacją (*cutoff-to-repair*) odnotowali Gyarmathy i Neuberger (2013). Jest to zapewne efekt rozwoju bardziej sprawnego programowania mowy, co wiąże się z (przynajmniej częściową) automatyzacją tego procesu na poziomie kodowania gramatycznego i morfofonologicznego (Levelt, 1989; Postma i Kolk, 1993; por. Hartsuiker i Moors, 2017). Według Gyarmathy i Neuberger (2013) dorośli mają dużo łatwiejszy dostęp leksykalny do wyrazów pomocniczych, gdyż używają ich jako „frazesów” (*clichés*). Objawia się to tym, że – w przypadku falstartów – faza edycji przed wyrazami pomocniczymi jest u dorosłych znacznie krótsza niż przed wyrazami treściowymi. Wyrazy pomocnicze nie ulegają też zniekształceniom dźwiękowym w wypowiedziach dorosłych użytkowników języka, co zdarza się w mówieniu małych dzieci,

---

118 Według Sabina i in. (1979) długość i częstość pauz niewypełnionych zmniejsza się najintensywniej w okresie przedszkolnym i od około 2 klasy szkoły podstawowej pozostaje stabilna aż do dorosłości, ale łączna długość pauz niewypełnionych jest najniższa w adolescencji.

które dopiero rozwijają fonologię, ale gramatycznie nie różnicują tych dwóch klas wyrazów (Wijnen, 1992; por. Dell, 1986).

Z drugiej strony, jak zauważają Sabin i in., (1979) od okresu adolescencji do dorosłości tempo mówienia nieznacznie spada, co interpretują oni jako wzrost refleksyjności i retorycznego wymiaru mowy. Uwaga ta pozwala odnieść opisane powyżej zmiany rozwojowe do ustalonego w rozdziale 3. związku kontroli procesu mówienia z funkcjami wykonawczymi. O ile taki związek wykazano w badaniach nad mówieniem osób dorosłych, niewiele do tej pory wykonano badań sprawdzających na ile postęp w monitorowaniu mowy można powiązać z dojrzewaniem funkcji wykonawczych. Zazwyczaj, jak już zauważono, zmiany związane z monitorowaniem i korygowaniem własnej mowy przypisuje się rozwojowi językowemu pod względem gramatycznym (Rispoli, 2003), fonologicznym (Sasisekaran i Weber-Fox, 2012) i semantycznym (Hanley i in., 2016). Pogląd ten wyrażają słowa Tarkowskiego (1992, s. 34) „zauważanie i korygowanie błędów w wypowiedziach własnych i cudzych jest przejawem kompetencji językowej. Dziecko jest poprawiane przez otoczenie (głównie rodziców), zanim zacznie poprawiać siebie. Czynić to może dopiero wówczas, gdy osiągnie odpowiedni poziom językowej sprawności systemowej”.

### **4.3 Rozwój kontroli procesu mówienia a rozwój funkcji wykonawczych**

Uwzględniając jednak związek monitorowania z kontrolą wykonawczą oraz fakt jej rozwoju aż do okresu wczesnej dorosłości, należałoby spodziewać się, że poziom sprawności funkcji wykonawczych będzie przekładał się na odmienną sprawność poznawczej kontroli procesu mówienia u osób w różnym wieku. Interesujące wyniki w tej kwestii przedstawia badanie Engelhardta i in. (2013). Wykazali oni, że wśród badanych adolescentów i dorosłych niższy poziom kontroli hamowania związany był z występowaniem większej liczby niepełności w postaci reperacji. Większość autonapraw (~60%) nie dotyczyła jednak korekty błędów leksykalnych, czy fonologicznych, a rewizji szyku zdania i wymiany jego składników (np. zamiana podmiotu z dopełnieniem)<sup>119</sup>. Wygląda więc na to, że niższe wyniki w teście hamowania reakcji, wiązały się z większą tendencją do impulsywnego produkowania wypowiedzi, które następnie w wyniku monitorowania były zatrzymywane i korygowane. Jest jeszcze inna możliwość, mianowicie osoby z wyższym poziomem hamowania rzadziej

---

<sup>119</sup> Przykładem takiej reperacji jest wypowiedź: *the girl ... the bicycle was ridden by the girl* (Engelhardt i in., 2013, s. 10).

zatrzymywały poprawne wypowiedzi, bardziej elastycznie tolerując formę językową, w jakiej wyrażali przekazywaną treść. W tym drugim przypadku należałoby jednak spodziewać się związków występowania reperacji z funkcją przełączania (giętkości poznawczej), czego nie wykazano w tym badaniu. Co ciekawe inne analizowane przez badaczy nie płynności nie przejawiały związku z funkcjami wykonawczymi. Niestety autorzy badania (Engelhardt i in., 2013) nie przedstawili wyników z podziałem na grupy wiekowe, nie wiadomo więc jakie różnice występowały między badanymi adolescentami i osobami dorosłymi i jakie znacznie dla reperacji (i innych nie płynności) mógł mieć związany z wiekiem poziom funkcji wykonawczych.

Badanie w grupie klinicznej dzieci jękających się<sup>120</sup> pokazało, że w porównaniu do swoich rówieśników (bez zaburzeń płynności mowy) wykazują one niższe wskaźniki hamowania, mierzone zadaniem Go/No-go (Eggers i in., 2013). Dzieci jękające się dawały więcej przedwczesnych odpowiedzi i ulegały fałszywym alarmom, nie wyhamowując swojej reakcji na sygnał „No-go”, a czas ich reakcji na fałszywe alarmy był krótszy. Ponadto w mniejszym stopniu były one w stanie dostosować swój styl reagowania po wystąpieniu błędnej reakcji. Podobnie, z dokonanego przez Anderson i Ofoe (2019) systematycznego przeglądu literatury wynika, że w wielu badaniach wykazano związki między funkcjami wykonawczymi a jękaniem. Po pierwsze, dzieci jękające się, szczególnie w wieku przedszkolnym, zazwyczaj osiągają niższe wyniki w zakresie kontroli hamowania niż ich niejękający się rówieśnicy, niezależnie czy wyniki uzyskiwane były na podstawie danych behawioralnych, czy oceny dokonywanej przez rodziców. Po drugie, wyniki badań nad związkiem jękania i poziomu pamięci roboczej nie są spójne. Niektóre badania pokazują, że nie ma różnicy między grupami klinicznymi a kontrolnymi, inne natomiast – skupiające się głównie na pamięci krótkotrwałej (powtarzanie pseudosłów lub cyfr) – zazwyczaj wykazują niższy poziom tej funkcji u dzieci jękających się w porównaniu do niejękających się rówieśników. Po trzecie, z niewielu przeprowadzonych jak dotąd badań wynika, że również pod względem giętkości poznawczej grupy dzieci jękających się wypadają gorzej na tle dzieci bez zaburzeń płynności mowy. Można więc ogólnie stwierdzić, że poziom funkcji wykonawczych pozwala do pewnego stopnia przewidywać ryzyko jękania się u dzieci. Warto jednak zaznaczyć, że w opublikowanej rok wcześniej metaanalizie przeprowadzonej przez tych samych badaczy (Ofoe i in., 2018) udało się potwierdzić związek jękania z zakresem pamięci krótkotrwałej oraz poziomem funkcji hamowania, ale mierzonej przez oceny

---

120 Główne objawy jękania odnotowane w badaniu polegały na powtarzaniu pojedynczych głosek lub sylab, wyrazów jednosylabowych, przedłużania artykulacji głosek lub blokowania („zacinań”) się na jakimś wyrazie (Eggers i in., 2013).

rodziców. Wskaźniki behawioralne hamowania okazały się nieróżnicujące dla grup klinicznych i kontrolnych.

Odwołanie się w tym miejscu do badań nad jękaniami w dzieciństwie jest o tyle zasadne, że część psycholingwistów wprost wiąże zjawisko jękania ze specyfiką pracy monitora, kontrolującego produkcję mowy. Mając więc przesłanki w postaci przypuszczalnych związków (a) między monitorowaniem a funkcjami wykonawczymi, (b) funkcjami wykonawczymi a jękaniami w dzieciństwie, (c) jękaniami a monitorowaniem produkcji mowy, zasadnym jest postawione pytanie o relacje między rozwojem funkcji wykonawczych w dzieciństwie i adolescencji, a rozwojowym progresem w monitorowaniu i kontroli mowy osiąganym wraz z wiekiem<sup>121</sup>. Poniżej zestawiono dwa popularne, kontrastujące ze sobą psycholingwistyczne podejścia do relacji między monitorowaniem produkcji mowy, a jękaniami.

1. Hipoteza ukrytych reperacji (*Covert Repair Hypothesis*). Według Postmy i Kolka (1993) za różne rodzaje nie płynności mówienia (blokady, przedłużenia artykulacji głosek, powtarzanie głosek lub sylab, przerywanie wyrazów, korygowanie błędów) może odpowiadać jeden wspólny mechanizm monitorowania wewnętrznego i dokonywania ukrytych reperacji. Zdaniem tych autorów nie płynności mówienia pojawiają się jako efekt uboczny wykrycia przez monitorowanie wewnętrzne błędu w programie mowy i uruchomienia procedury reperacji tego błędu, przed jego zewnętrzną artykulacją. Postma i Kolk (1993) uważają, że nie ma jakościowej różnicy między zaburzeniem płynności mowy w jękaniu, a nie płynnością występującą w mowie typowych użytkowników języka, gdyż wszystkie nie płynności powstają na bazie tego samego mechanizmu ukrytych reperacji. Różnica między obiema grupami ludzi sprowadza się natomiast do nasilenia tej nie płynności w jękaniu, a więc do różnic ilościowych, które badacze wiążą z komplikacjami na poziomie kodowania fonologicznego. Ich zdaniem osoby jękaące się mają zaburzoną zdolność do generowania programów mowy wolnych od usterek (np. w wyniku zbyt szybkiego

---

121 Należy w tym miejscu zaznaczyć, że niniejszej pracy rozważana jest relacja między monitorowaniem produkcji mowy a funkcjami wykonawczymi tzw. „zimnymi”. Ich rozwój na przestrzeni pierwszych dwóch dekad życia mógłby wyjaśniać postępującą sprawność w kontroli procesu mówienia, osiąganą stopniowo przez dzieci i młodzież aż do okresu dorosłości. Zależność ta mogłaby też tłumaczyć relatywnie niższy poziom funkcji wykonawczych u osób jękaących się, przy założeniu, że dla występowania symptomów jękania kluczowe znaczenie mają mechanizmy monitorowania i kontroli mówienia (powiązane z funkcjami wykonawczymi).

Niewykluczone jednak, że dla jękania, szczególnie tzw. nierozwojowego (por. Tarkowski, 1992), które ma charakter kliniczny ważniejsze są czynniki związane z samoregulacją emocjonalną i tzw. „zimnymi” funkcjami wykonawczymi (Putko, 2008; Diamond, 2013). Na przykład w jednym z cytowanych przez Anderson i Ofoe (2019) badań poziom wytężonej kontroli zachowania (*effortful control*) negatywnie korelował z nasileniem objawów jękania u dzieci (Kraft i in., 2014, za: Anderson i Ofoe, 2019).

generowania fonologicznej reprezentacji wypowiedzi) przez co popełniają więcej błędów w kodowaniu fonologicznym, a to z kolei skutkuje nasileniem sygnałów alarmowych dla monitora i ukrytymi reperacjami tych błędów. W zależności od tego w którym miejscu struktury wyrazu występuje błąd oraz kiedy (w którym miejscu artykułowanego wyrazu) nastąpi interupcja nie płynności przybierają różną postać – od powtórzeń pierwszych głosek, po powtórzenia całych wyrazów, blokady, przeciągnięcia artykulacji itp. Hipoteza ukrytych reperacji dobrze tłumaczy, dlaczego zaburzenia płynności mowy (u osób dorosłych) dotyczą częściej wyrazów treściowych niż funkcyjnych (Hartsuiker in., 2005), nie uzyskała jednak pełnego potwierdzenia w badaniach empirycznych. Po pierwsze trudno jest zweryfikować założenie o tym, że za każdym zająknięciem stoi błąd w programowaniu, który jest wykrywany przez monitorowanie wewnętrzne i poprawiany w sposób wywołujący zaburzenia płynności. Błędy takie są z definicji nieobserwowalne zewnętrznie i jedynym ich przejawem są ukryte reperacje. Założenie to wymagałoby przynajmniej wykazania, że kodowanie fonologiczne u osób jękających się nie przebiega optymalnie lub nie jest w pełni dojrzałe, czego badania nie potwierdzają (Brocklehurst, 2008).

2. Hipoteza błędnego koła (*Vicious Cycle Hypothesis*). O ile koncepcja Postmy i Kolka (1993) postuluje prawidłowe działanie mechanizmów monitorowania mowy (przy zaburzonym kodowaniu fonologicznym) u osób jękających się, konkurencyjne podejście prezentowane przez Vasić i Wijnena (2005) zakłada, że przyczyną jękania się jest wadliwie działający monitor mowy. Według tej hipotezy osoby jękające cechuje nadmierna czujność w monitorowaniu pewnych aspektów własnej mowy, gdyż przyjmują oni zbyt wysokie kryterium poprawności wypowiedzi. Z tego powodu proces monitorowania za bardzo koncentruje się na nadchodzących nie płynnościach lub problemach z planowaniem i często przerywa artykulację, co paradoksalnie nasila nie płynności zamiast im zapobiegać (Hartsuiker i in., 2005). Vasić i Wijnen (2005) wychodzą z założenia, że monitorowanie jest powiązane z procesami uwagowymi (por. Levelt, 1989; Levelt i in., 1999), przez co może być selektywnie ukierunkowane na inne aspekty generowanych dźwięków mowy (wysokość głosu, głośność, poprawność stylistyczna wypowiedzi, dokładność artykulacji, poprawność językowa itd.). Zasoby poznawcze, jakie są potrzebne do monitorowania mowy mogą być w różny sposób lokowane w zależności od tego, (1) jak dużo tych zasobów jest zużywane (*effort*) oraz (2) na jaki aspekt strumienia mowy są kierowane (*focus*). Trzeci parametr monitorowania określany jest mianem (3) „progu” (*threshold*) i

odnosi się do kryteriów, które dane wyjściowe muszą spełniać, aby były akceptowane jako poprawne (pod względem absolutnych zasad języka oraz indywidualnych standardów osoby mówiącej). Hipoteza błędnego koła zakłada, że te trzy parametry monitora mogą mieć niewłaściwe „ustawienia” u osób jękających się, co znaczy, że

- a) więcej niż potrzeba wysiłku (*effort*) jest inwestowane w monitorowanie procesu mówienia;
- b) monitor skupia się (*focus*) czasowych fluktuacjach i nieciągłościach mówienia (które są normalne w spontanicznej mowie);
- c) kryteria progowe (*threshold*) akceptowalnego efektu produkcji mowy są zbyt wysokie, przez co wzbudzają fałszywy alarm błędu nawet przy niewielkich (i nieuniknionych w spontanicznej mowie) czasowych fluktuacjach i nieciągłościach, które zaczynają być postrzegane jako niepożądane zakłócenia płynności mowy.

W ten sposób, paradoksalnie, jękanie pojawia się jako efekt uboczny pracy monitora nastawionej na przeciwdziałanie niepłynności mowy (tzw. błędne koło). Hipoteza błędnego koła nie wyklucza tego, że część niepłynności jest skutkiem wykrycia przez monitor błędów w programowaniu i prób ich reperacji. W przeciwieństwie do hipotezy ukrytych reperacji nie zakłada jednak, że podstawą każdej ukrytej naprawy jest błąd (Bernstein Ratner i Wijnen, 2007). Proponuje raczej, że w części przypadków monitor reaguje na zjawiska, które w sposób nieunikniony towarzyszą planowaniu mowy, a szczególnie na te, które wiążą się z jakąś (pozorną) nieciągłością. W ten sposób proces monitorowania staje się dezadaptacyjny i zaburza płynność mowy, co odbierane jest jako jękanie. Hipoteza błędnego koła dobrze wyjaśnia znane zjawiska znoszenia symptomów jękania przy zewnętrznych źródłach koordynacji czasowej mówienia (*timing*) takich jak śpiewanie oraz przy modyfikacji percepcji własnej mowy w mówieniu z echem oraz opóźnionym słuchowym sprzężeniem zwrotnym (*delayed auditory feedback*). Ponadto Vasić i Wijnen (2005) dowiedli, że wykonywanie dodatkowego zadania (gra PONG angażująca funkcje wzrokowo-motoryczne) przez osoby jękające się podczas mówienia zmniejsza stopień niepłynności, gdyż angażuje zasoby poznawcze, które typowo (zgodnie z hipoteza) inwestowane są w intensywne monitorowanie mowy (por. Hartsuiker i in., 2005). Dodatkowo badacze ci wykazali, że przekierowanie uwagi na inne niż zwyczajowo monitorowane aspekty mówienia (osoby badane miały naciskać przycisk za każdym razem, kiedy wykryją we własnej mowie zaimbek „to”) także prowadzi do redukcji symptomów niepłynności u jękających się uczestników badania.

Hipoteza błędnego koła, przez swoje odwołanie do zaangażowania zasobów poznawczych w proces monitorowania jest spójna z wynikami wiążącymi jąkanie z wadliwym działaniem funkcji wykonawczych. Według Paphiti i Eggersa (2022), którzy wykazali, że jākające się dzieci w wieku szkolnym (7-10 lat) prezentują niższy poziom giętkości poznawczej niż ich płynnie mówiący rówieśnicy, uważają że uzyskane przez nich wyniki zgadzają się z hipotezą błędnego koła. Autorzy tego badania twierdzą, że niższy poziom elastyczności może oznaczać, że podczas mówienia dzieci jākające się mogą mieć większą trudność w odwróceniu uwagi od popełnianych błędów oraz wtórnie wpływać na inne symptomy jąkania.

Z powyższych ustaleń wynika więc, że nie płynności mowy obserwowane w jąkaniu u dzieci można do pewnego stopnia wiązać z procesem monitorowania mowy (prawidłowo reagującym na nasilone trudności w programowaniu wypowiedzi lub nadmiernie restrykcyjnym względem typowych potknięć w mówieniu) oraz z niższym niż typowy poziomem funkcji wykonawczych. Niestety nadal niewiele wiadomo na temat tego jak poziom funkcji wykonawczych różnicuje umiejętności monitorowania i kontroli mowy u zdrowo rozwijających się dzieci i młodzieży. W cytowanym wcześniej badaniu Hanley i in. (2016) kontrolnie zbadano zakres pamięci roboczej u uczestników (powtarzanie cyfr wprost i wstak z testu WISC-IV), ale nie wykazano związku tego parametru z malejącą wraz z wiekiem tendencją do popełniania błędów leksykalnych i wzrastającą sprawnością ich reperacji. Niestety badacze – pomimo tego iż odwołują się do teorii monitorowania konfliktu Nozari i in. (2011) – nie uwzględnili w badaniu funkcji hamowania, która jest kluczowa dla wykrywania błędów przez system kontrolny zlokalizowany w przednim zakręcie kory obręczy (ACC) (Nozari, 2020; por. Fąfrowicz i Marek, 2008). W związku więc z brakiem konkluzywnych badań na temat relacji między rozwojem funkcji wykonawczych, a postęпами monitorowania i kontroli procesu mówienia w późnym dzieciństwie i adolescencji przeprowadzono badanie, które opisane zostało w dalszej części rozprawy.

## 5 Badanie własne

### 5.1 Wprowadzenie

W poprzedniej części pracy przedstawione zostało poznawcze podejście do zagadnienia procesu mówienia i jego kontroli oraz ich tło rozwojowe. Na podstawie omówionej literatury wysnuć można wniosek, że zdolność do monitorowania i kontroli produkcji mowy jest powiązana z funkcjami wykonawczymi, odpowiedzialnymi za kontrolę czynności celowych. Podobnie też, jak funkcje wykonawcze, kontrola produkcji mowy rozwija się i doskonali wraz z wiekiem, osiągając szczytową formę swojego funkcjonowania w okresie wczesnej dorosłości.

W literaturze psycholingwistycznej rozwój skutecznego programowania i korygowania własnych wypowiedzi rozpatrywany jest głównie na tle nabywania języka. Dużo mniej uwagi poświęcono paraleli rozwoju funkcji wykonawczych i rozwojowym zmianom w kontroli procesu mówienia. Ze względu na niedostatek prac empirycznych z tego zakresu podjęto decyzję, aby zaprojektować badanie, które pozwoli odpowiedzieć na pytanie o to w jaki sposób (jeśli w ogóle) kontrola procesu mówienia zmienia się wraz z wiekiem od okresu późnego dzieciństwa do dorosłości, a także w jaki sposób kontrola procesu mówienia wiąże się z poziomem funkcji wykonawczych.

Zanim przedstawione zostaną szczegółowo problemy i hipotezy badawcze (zob. sekcja 5.2.), przybliżyć należy kluczowe aspekty kontroli procesu mówienia (oraz ich behawioralne przejawy), które wybrano do badania przeprowadzonego w ramach niniejszego projektu. Jak wspomniano w rozdziale 3. monitorowanie produkcji mowy jest procesem dyskretnym, który przebiega niezauważalnie, o ile proces produkcji mowy nie napotyka na poważniejsze trudności. W takiej sytuacji, kiedy tekst wypowiedzi tworzony jest płynnie, kontrola procesu mówienia pozostaje w tle. Kiedy jednak w procesie mówienia na którymś z poziomów przetwarzania wystąpi problem (potencjalne trudności mogą dotyczyć selekcji leksykalnej i kodowania gramatycznego, kodowania morfo-fonologicznego, tworzenia planu fonetycznego lub artykulacji) może to skutkować różnymi zakłóceniami płynności w postaci zawahania, przerwy w mówieniu, falstartu, powtórzenia, reorganizacji wypowiedzi itp. Wystąpienie tych zjawisk jest nie tylko przejawem trudności w programowaniu mowy, ale też – pośrednio – efektem działania mechanizmów kontrolnych, które utrzymują tok mówienia i przywracają płynność mowy zaburzoną, np. pojawieniem się błędu w programie mowy<sup>122</sup>.

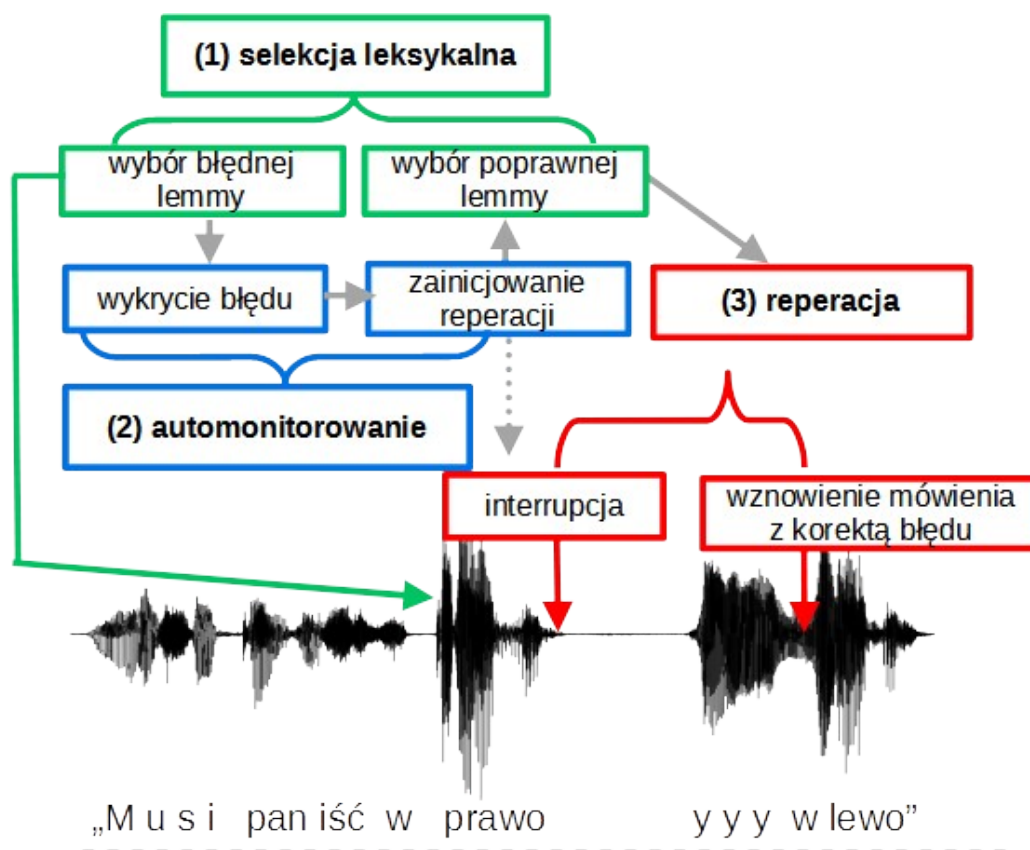
---

<sup>122</sup> Choć generalnie odróżnia się niepłynności (*dysfluencies*) od błędów mowy (*speech errors*) (Bock, 1996), niektórzy badacze traktują wszystkie te zjawiska jako „błędy mowy” w szerokim rozumieniu tego pojęcia (Cutler i Fay, 1978).

Szczególnym przypadkiem błędów mowy są przejęzyczenia (błędy mowy w wąskim znaczeniu), czyli wypowiedzi, w których generowany tekst odbiega od zamierzonego, nawet jeśli nie są łamane zasady poprawności językowej. Dzięki monitorowaniu procesu mówienia wiele błędów udaje się wykryć i poprawić, często zanim dojdzie do ich faktycznego wyartykułowania (Levelt, 1983; Postma i Kolk, 1993). W związku z tym, że wyjściowy dla niniejszej pracy model mówienia zaproponowany przez Levelta przyjmuje za podstawową jednostkę analizy generowanie pojedynczych wyrazów (Levelt i in., 1999; Levelt, 1999a); i/lub zdań (Levelt, 1999b; Levelt, 2000) przejęzyczenia dokonywane na tym materiale (w obrębie zdań lub pojedynczych wyrazów) uznać można za przypadki modelowe dla badania kontroli procesu mówienia.

### Rysunek 8

*Wybrane aspekty kontroli procesu mówienia na przykładzie przejęzyczenia i jego reperacji*



W rozdziale 3. przedstawiono przykład typowego przejęzyczenia i jego autokorektę, pozwalające na wnioskowanie na temat działających w takiej sytuacji mechanizmów kontrolnych mowy. Dla przypomnienia, kanoniczny przykład przejęzyczenia zilustrowany na rysunku 8 składa się z:

- (1) reparandum, czyli błędnego wyrazu, frazy lub ich części;
- (2) momentu interrupcji, czyli momentu w którym przerwana zostaje błędna wypowiedź;
- (3) fazy edycji, czyli odcinka czasu pomiędzy przerwaniem artykulacji, a jej wznowieniem (*cutoff-to-repair*)
- (4) reperatury, czyli wznowionej wypowiedzi, która realizuje korektę popełnionego błędu.

Podany modelowy przykład przejęzyczenia stanowi punkt odniesienia dla przeprowadzonego i opisanego w dalszej części badania mechanizmów kontroli procesu mówienia. Na potrzeby operacjonalizacji i pomiaru zmiennych wyróżniono trzy kluczowe elementy powstawania przejęzyczenia (na poziomie leksykalnym) i jego korekty, które uwydatniają trzy różne aspekty kontrolne w produkcji mowy:

1. Selekcja leksykalna – jeden z etapów programowania mowy, odpowiedzialny za wybór właściwej lemmy i odpowiadającej jej słowoformy (por. rozdział 2). Przejęzyczenie polegające na zamianie słów może być efektem dystrakcji, która zakłóca wybór właściwej lemmy na rzecz konkurencyjnej jednostki poprzez aktywowanie innych pojęć leksykalnych, semantycznie powiązanych z wyrazem, który ma zostać wypowiedziany. Według modelu Levelta i in. (1999; Roelofs, 2020) zadaniem działających na poziomie leksykonu mechanizmów weryfikacyjnych jest niedopuszczanie do realizacji lemmy, która została aktywowana w wyniku procesów innych niż realizacja intencji komunikacyjnej. Problem badawczy dotyczący tego aspektu kontroli procesu mówienia wiąże się z pytaniem o to, czy wraz z wiekiem zmienia się skuteczność selekcji leksykalnej i jej odporność na dystrakcję oraz to, w jakim stopniu poziom kontroli selekcji leksykalnej jest związany z poziomem kontroli wykonawczej u osób w różnym wieku (zob. 5.1.1.).

2. Automonitorowanie tworzonej wypowiedzi – jest procesem służącym sprawdzaniu generowanego tekstu pod względem jego poprawności i zgodności z intencją. Dzięki monitorowaniu własnej mowy możliwa jest detekcja popełnionej pomyłki i przerwanie błędnej wypowiedzi oraz zainicjowanie działań korekcyjnych. Problem badawczy dotyczący tego procesu wiąże się z pytaniem o to, czy wraz z wiekiem zmienia się efektywność automonitorowania procesu mówienia oraz w jakim stopniu powiązana jest ona z poziomem kontroli wykonawczej u osób w różnym wieku (zob. 5.1.2.).

3. Interrupcja błędnej wypowiedzi i reperacja – stanowią podstawowy sposób reagowania na popełnione przejęzyczenie. Po wykryciu błędu osoba mówiąca staje przed zadaniem naprawienia go, co pociąga za sobą wymaganie przerwania wypowiedzi zawierającej reparandum oraz zaplanowanie i realizację korekty. Problem badawczy dotyczący tego procesu wiąże się z pytaniem o to, czy wraz z wiekiem zmienia się sposób dokonywania reperacji (interrupcja i korekta) błędnych wypowiedzi oraz w jakim stopniu różnice w dokonywaniu korekt przez ludzi w różnym wieku są związane z poziomem kontroli wykonawczej (zob. 5.1.3.).

Każdy z wyróżnionych powyżej aspektów kontroli procesu mówienia może nie tylko być badany w kontekście rzeczywistych przejęzyczeń, ale także symulowany w odrębnych zadaniach laboratoryjnych, które opracowane zostały na potrzeby niniejszej rozprawy: Nazywanie obrazków z dystrakcją (paradygmat badawczy: *picture-word interference paradigm*), Monitorowanie sylab, Powtarzanie zdań (paradygmat badawczy: *repair elicitation paradigm*). Kwestia ta omówiona została szczerzej w kolejnych sekcjach (5.1.1.-5.1.3.).

### **5.1.1 Odporność selekcji leksykalnej na dystrakcję i efekt interferencji semantycznej**

Nie jest do końca jasne w jaki sposób dochodzi do różnego rodzaju przejęzyczeń. Na przykład model produkcji mowy Della (1986) traktuje powstawanie przejęzyczeń jako konsekwencję pewnej elastyczności, która cechuje system produkcji mowy, pozwalając na kreatywne wykorzystanie potencjału systemu językowego. W tym modelu powstawanie przejęzyczeń ujmowane jest jako proces probabilistyczny, pozostający pod wpływem czynników, takich jak częstość używania słów, podobieństwo semantyczne (lub fonologiczne) itp., które wpływają na poziom aktywacji poszczególnych jednostek w sieci leksykalnej. Choć poszczególne modele różnią się co do szczegółów opisu przebiegu procesu mówienia (zob. rozdział 2.), ogólnie zgadzają się, że w zależności od momentu programowania, w którym doszło do selekcji błędnego elementu mogą działać inne czynniki facylitujące lub ograniczające powstawanie przejęzyczeń (Bręński, 2015). Część błędów w selekcji leksykalnej może być np. efektem dystrakcji na jaką użytkownicy języka są wystawieni podczas mówienia. Dystrakcja ta może pochodzić z samego systemu produkcji mowy (np. z leksykonu umysłowego), kiedy w procesie selekcji leksykalnej konkurują ze sobą dwie jednostki o bardzo wysokim potencjale aktywacji w danym kontekście. W podanym na rysunku 8. przykładzie alternatywą dla wyrażenia „w lewo” jest wyrażenie „w prawo”, przy czym

kryterium poprawności użycia któregoś z tych wyrażen ma charakter względny, a nie absolutny. Z tego powodu pomyłka polegająca na niewłaściwym nazwaniu wskazywanego kierunku ma stosunkowo wysokie prawdopodobieństwo wystąpienia, gdyż w zależności od przyjętej perspektywy, każde z określeń kierunku może być poprawne lub błędne (w zależności od układu odniesienia).

Źródłem dystrakcji mogą być również czynniki zewnętrzne względem czynności mówienia (np. przeczytany napis w gazecie lub na szyldzie reklamowym, treść słyszanej w tle rozmowy lub audycji, a także treść własnych myśli niezwiązanych z tworzoną wypowiedzią itp.), skutkujące popełnieniem błędu (Harley, 1984; 2001) lub przynajmniej zakłóceniem płynności wypowiedzi. Według Clarka (1996) najczęstszym źródłem zakłóceń podczas rozmowy jest nakładająca się mowa (*overlapping speech*) innych osób. Zjawisko to, czyli wpływ zewnętrznej dystrakcji na proces mówienia można symulować w kontrolowany sposób w procedurze nazywania obrazków z interferencją (*picture-word interference paradigm*), w której osoby badane podczas wypowiadania nazw prezentowanych obrazków jednocześnie mają ignorować prezentowane im równolegle wyrazy-dystraktory. Jeśli dystrakcja ma charakter semantyczny (nazwa obrazka powiązana jest znaczeniowo z towarzyszącym obrazkowi dystraktorem) obserwuje się tzw. efekt semantycznego hamowania (*semantic inhibition effect*) zwany też efektem interferencji semantycznej (*semantic interference effect*), polegający na opóźnieniu czasu reakcji werbalnej (podanie nazwy obrazka) w próbach z dystraktorem w porównaniu do prób bez dystrakcji<sup>123</sup> (bardziej szczegółowo o zastosowanej w niniejszym projekcie procedurze nazywania obrazków z interferencją traktuje sekcja 5.4.2.1.1.).

Badania pokazują, że średni czas reakcji w zadaniu nazywania obrazków skraca się wraz z wiekiem i wraz z wiekiem maleje też (choć nieznacznie) interferujący wpływ dystrakcji semantycznej (Rosinski, 1977). Jednocześnie czas reakcji, badany w grupie osób dorosłych, dodatnio koreluje z czasem hamowania w zadaniu sygnału stop, mierzącym poziomem funkcji hamowania (Shao i in., 2012), także kiedy kontrolowana jest zmienna wieku (Shao i in., 2013). Jeśli przyjąć, że zadanie nazywania obrazków z interferencją mobilizuje mechanizmy kontrolne „góra-dół” (np. mechanizm weryfikacji aktywowanych lemm w odniesieniu do realizowanego pojęcia leksykalnego) i symuluje naturalne procesy

123 Generalnie prezentowanie innych wyrazów podczas nazywania obrazków ma charakter dystrakcyjny, jednak w zależności od relacji czasowej między wyświetleniem obrazka a zaprezentowaniem dystrakcyjnego wyrazu, efekt interferencji będzie miał różną wielkość. Co więcej, w przypadku wyrazów przedstawiających podobieństwo fonologiczne względem nazwy obrazka może dojść do efektu facylitacji, czyli przyspieszenia czasu nazywania. Facylitacja jest obserwowana, kiedy obrazek wyświetlany jest jako pierwszy bodziec, a nazwa podobnego fonologicznie dystraktora pojawia się kilkaset milisekund później (SOA <50-300 ms>) Największy efekt interferencji semantycznej obserwuje się natomiast przy relacji czasowej między bodźcami SOA <-100;+100ms> (Schriefers i in., 1990; Levelt, 1993).

występujące w spontanicznej mowie, to wyniki tego zadania można wykorzystać w próbie odpowiedzi na pytania o to, czy wraz z wiekiem zmienia się skuteczność selekcji leksykalnej i jej odporność na dystrakcję oraz w jakim stopniu poziom kontroli selekcji leksykalnej jest związany z poziomem kontroli wykonawczej u osób w różnym wieku. Szczegółowe pytania badawcze i hipotezy związane z tym problemem przedstawione zostaną w podrozdziale 5.2. Szczegółowy opis zadania wybranego do zbadania kontroli selekcji leksykalnej znaleźć można w sekcji 5.4.2.1.1.

### **5.1.2 Automonitorowanie procesu mówienia**

Monitorowanie procesu mówienia (automonitorowanie) odpowiada za detekcję przejęzyczeń i innych błędów lub nieścisłości w realizacji intencji komunikacyjnej. Przywołana w rozdziale 3. główna zasada przerywania mowy (*main interruption rule*), zakładała, że artykulacja przerywana jest – z pewnymi wyjątkami – niezwłocznie po wykryciu błędu. Tym samym moment przerywania mówienia mógłby być traktowany jako behawioralny wskaźnik detekcji błędu, a czas pomiędzy początkiem reparandum, a przzerwaniem mówienia (*error-to-cutoff*) wskazywałby czas potrzebny procesowi monitorowania na wykrycie pomyłki i reakcję. Jak już wcześniej wspomniano, badania pokazują jednak, że choć detekcja błędu i planowanie reperacji są procesami sekwencyjnymi, ich manifestacja w postaci następujących po sobie momentów: zatrzymanie artykulacji, pauza (faza edycji) i wznowienie mówienia, nie jest konieczna. Część błędów, która wykrywana jest przez monitorowanie wewnętrzne nie powoduje natychmiastowej interupcji, a planowanie reperacji odbywa się jeszcze w trakcie artykulacji reparandum bez wyraźnej fazy edycji (Blackmer i Mitton, 1991; Hartsuiker i Kolk, 2001; Seyfeddinipur, i in., 2008). Tym samym moment interupcji nie jest adekwatnym wskaźnikiem ani momentu detekcji błędu (wykryty wewnętrznie błąd nie musi wywoływać natychmiastowej interupcji) ani nawet tego, czy do detekcji błędu doszło (błąd może zostać wykryty i naprawiony bez interupcji reparandum). Chcąc zatem badać skuteczność monitorowania wewnętrznego oraz jego czasowe charakterystyki należy oprzeć się na innych behawioralnych markerach.

Niezawodnym wskaźnikiem skutecznej detekcji błędu nie może być także sam fakt dokonania reperacji. O ile w przypadku popełnienia błędu i jego korekty reperacja jest dowodem na to, że błąd został wykryty, o tyle przypadki błędów pozostawionych bez naprawy nie muszą jednoznacznie świadczyć o braku wykrycia ich przez proces monitorowania. Część błędów mimo ich detekcji, może być tolerowana, jako nieistotne zakłócenie nie wymagające reperacji. Innym jeszcze problemem jest klasyfikowanie

falstartów, czy powtórzeń jako przejawów reperacji ukrytych błędów. Nie wszyscy badacze zgadzają się, że są to niezawodne wskaźniki reperacji błędów wykrytych przez monitorowanie wewnętrzne (por. Postma, 2000).

Niektórzy badacze angażując osoby badane w zadania generujące przejęzyczenia, np. łamańce językowe, w procedurze badawczej wymagają sygnalizowania popełnienia błędu w postaci naciśnięcia klawisza aparatury badawczej (Hartsuiker, Kolk i Martensen, 2005). Podobnie jednak, jak w przypadku błędów bez reperacji, nie można mieć pewności, że każdy wykryty błąd w mówieniu został zasygnalizowany przez naciśnięcie klawisza (jest to być może łatwiejsze niż dokonanie reperacji, szczególnie pod presją czasu, ale jednak funkcjonalnie odległe od typowej reakcji na popełnienie błędu). Nie sposób wykluczyć z tej procedury wypadków, kiedy błąd został wykryty, ale osoba badana nie zareagowała na niego poprzez naciśnięcie odpowiedniego klawisza<sup>124</sup>.

Jak pokazują badania psycholingwistyczne (np. Wheeldon i Levelt, 1995; Wheeldon i Morgan, 2003; Schiller, 2005; Özdemir i in. 2007) monitorowanie procesu mówienia może być badane poza kontekstem przejęzyczeń. Zgodnie z modelem Levelta i in. (1999) monitorowanie wewnętrzne ma dostęp do efektów kodowania morfo-fonologicznego (wyrazów fonologicznych), które może sprawdzać – za pośrednictwem systemu odbioru mowy – pod kątem ich odpowiedniości względem pojęć leksykalnych (aktywowanych w procesie konceptualizacji). Jak pisze Roelofs (2020a, s. 3) „monitor porównuje wybrane reprezentacje produkcji i rozumienia<sup>125</sup>” (tzn. reprezentacje aktywowane do stworzenia wypowiedzi są porównywane z reprezentacjami aktywowanymi w systemie rozumienia mowy). Porównywanie to nie musi dotyczyć wyłącznie poprawności tych reprezentacji, ale też ich struktury. Według metodologii zaproponowanej przez Wheeldon i Levelta (1995), polecając osobom badanym wykrywanie wskazanych elementów subleksykalnych (np. fonemów, sylab itp.) w „wypowiadanych” (generowanych bezgłośnie lub na głos) wyrazach, można mieć wgląd w przebieg procesu monitorowania. W takim zadaniu monitorując wewnętrznie (lub zewnętrznie) generowane wyrazy (np. nazwy obrazków lub tłumaczenia wyrazów podanych w innym języku) wskazuje się poprzez naciśnięcie odpowiedniego klawisza (odp. TAK/NIE), czy w strukturze danego wyrazu znajduje się dany element (fonem, sylaba, rym). Do pewnego stopnia technika ta jest więc podobna do zadań, w których sygnalizuje się popełnienie błędu w planie wewnętrznym wypowiedzi. Tu również oczekuje

---

124 Z tego względu być może bardziej wiarygodnymi wskaźnikami niż behawioralne, wprost związane z mówieniem, byłyby wskaźniki na poziomie aktywności neuronalnej powiązane z wykrywaniem błędów. Na przykład Nozari (2020) jako wskaźnik wykrywania błędów przez monitorowanie wewnętrzne traktuje negatywną polaryzację (ERN) w zapisie EEG (paradygmat potencjałów wywołanych), która sygnalizuje aktywność przedniego zakrętu obręczy (Fąfrowicz i Marek, 2008).

125 Org. *the monitor compares selected production and comprehension representations* (Roelofs, 2020a, s. 3).

się, aby efekty monitorowania zostały potwierdzone przez wybór odpowiedniego klawisza (reakcja motoryczna). Różnica w sposobie udzielania odpowiedzi – poza tym, że zadanie monitorowania sylab lub fonemów nie dotyczy wykrywania błędów – polega na tym, że odpowiedzi mogą być odniesione do obiektywnego kryterium poprawności: odpowiedź twierdząca jest poprawna tylko wtedy, kiedy dany element rzeczywiście występuje w przetwarzanym wyrazie, a odpowiedź przecząca jest poprawna kiedy przetwarzany wyraz nie zawiera danego elementu<sup>126</sup>. Występuje tu więc reakcja motoryczna z wyborem („tak” vs. „nie”), zamiast reakcji motorycznej prostej sygnalizującej popełnienie błędu, której poprawności nie da się zweryfikować w generowaniu mowy bezgłosnej.

Chociaż metodologia badania monitorowania wewnętrznego, zaproponowana przez Wheeldon i Levelta (1995) nie służy wykrywaniu przejęzyczeń, to daje efekty zbieżne z danymi na temat monitorowania błędów w kodowaniu fonologicznym. Wykazali to Nootboom i Quené (2019), którzy w zadaniu łamańców językowych (*tongue twisters*) manipulowali miejscem wystąpienia przejęzyczenia w wypowiedzianym wyrazie. Okazało się, że w zależności od tego, czy błąd występował w nagłosie czy śródgłosie, czas między początkiem błędnego wyrazu, a interupcją był różny (dla wyrazów z błędem w śródgłosie dłuższy niż dla błędów w nagłosie). Podobnie, w badaniu Wheeldon i Levelta (1995) rozpoznawanie poszczególnych fonemów w generowanych bezgłosnie wyrazach miało charakter przyrostowy, a czas reakcji (potwierdzenie obecności danego fonemu w wyrazie) wydłużał się wraz z występowaniem fonemu docelowego na dalszych pozycjach wyrazu. Zestawiając wyniki obu badań Nootboom i Quené (2019) uważają, że monitorowanie wewnętrzne w celu wykrywania błędów oraz monitorowanie wewnętrzne w celu wykrywania fonemów w strukturze wyrazów są procesami porównywalnymi. Można więc przyjąć, że zadanie monitorowania fonemów (lub innych elementów subleksykalnych) symuluje warunki w których ujawnia się działanie automonitorowania mowy, a wyniki uzyskane w zadaniu tego typu odzwierciedlają ogólną sprawność monitorowania własnej mowy.

Biorąc pod uwagę problem rozwoju automonitorowania warto przypomnieć, że Sasisekaran i Weber-Fox (2012) wykazały, iż monitorowanie wewnętrzne wypowiedzi (bezgłosne nazywanie obrazków) w poszukiwaniu określonych elementów subleksykalnych zwiększa swoją sprawność w toku rozwoju dziecka oraz, że monitorowanie pojedynczych fonemów jest dla dzieci trudniejsze niż monitorowanie rymu (monitorowanie fonemów zajmuje więcej czasu). Ponadto we wspomnianym badaniu Wheeldon i Levelta (1995)

---

126 Słabą stroną tego paradygmatu badawczego jest to, iż poziom poprawności udzielanych odpowiedzi odzwierciedla nie tylko sprawność monitorowania, ale może być częściowo efektem poziomu opanowania samego sposobu udzielania odpowiedzi, który wymaga przypisania danego klawisza do danego typu odpowiedzi (TAK / NIE).

dodatkowe zaangażowanie osób badanych w głośne odliczanie podczas zadania monitorowania fonemów skutkowało niewielkim spadkiem poprawności rozpoznań fonemów oraz opóźnieniem czasu reakcji, względem średniej w grupie osób badanych bez zadania odliczania. Może to sugerować, że zasoby, z których korzysta monitorowanie są współdzielone z innymi procesami poznawczymi i mogą być ograniczone np. zakresem pamięci roboczej.

W związku z powyższym, można oczekiwać, że korzystając z zadania monitorowania elementów subleksykalnych (np. sylab) uda się odpowiedzieć na pytanie o to, czy wraz wiekiem zmienia się efektywność monitorowania procesu mówienia oraz w jakim stopniu powiązana jest ona z poziomem kontroli wykonawczej u osób w różnym wieku. Szczegółowe pytania badawcze i hipotezy związane z tym problemem przedstawione zostaną w sekcji 5.2., natomiast dokładny opis zastosowanego zadania badawczego znajduje się w punkcie 5.4.2.1.2.

### **5.1.3 Interrupcja mówienia i sposoby dokonywania reperacji**

Analiza reperacji jest nieodzownym elementem badania kontroli procesu mówienia w kontekście spontanicznych oraz eksperymentalnie indukowanych przejęzyczeń (Levelt, 1983; 1989; Nootboom, 1980; Nootboom i Quené, 2019; Postma i Oomen, 2005). Jednocześnie niewiele uwagi poświęcono do tej pory różnicom indywidualnym oraz zmienności rozwojowej w sposobie dokonywania reperacji przez typowych użytkowników języka. Nie wiadomo na przykład na ile szybkość dokonywania interrupcji może być funkcją zdolności hamowania reakcji, ani czy wraz z wiekiem zmienia się sposób korygowania błędów poprzez proporcjonalny wzrost reperacji ukrytych, które wymagają zaangażowania większych zasobów poznawczych (por. Hartsuiker i in., 2005).

W błędach takich, jak w przykładzie zilustrowanym na rysunku 8 po reparandum następuje interrupcja, a po czasie pewnego zawieszenia (*suspension*) wznowienie artykulacji i korekta w postaci zastąpienia błędnej lemmy „prawo” prawidłową „lewo”. Jak zauważają van Wijk i Kempen (1987) w przypadku substytucji lemmy (*lemma substitution*) nie dochodzi do rewizji struktury gramatycznej zdania, a reperacja polega na odtworzeniu części wypowiedzi (od początku zdania, frazy lub przynajmniej od miejsca sprzed wystąpienia błędu, por. Levelt, 1989) według zaplanowanego wcześniej szyku. Inaczej jest w przypadku tzw. przeformułowania (*reformulation*), które wymaga rewizji składniowej struktury wypowiedzi (Postma, 2000) i jest stosowane wtedy, gdy dla wyrażenia treści konieczna jest zmiana szyku wyrazów, np. przy zamianie strony czynnej zdania na stronę bierną.

Niezależnie od tego, jakie zmiany wprowadza korekta, dokonanie reperacji popełnionego błędu wymaga interupcji w reparandum. Główna zasada przerywania mówienia (*MIR*) definiuje interupcję jako natychmiastową (i poniekąd automatyczną) reakcję na wykrycie błędu (Levelt, 1989; Nooteboom, 1980). Po niej dopiero (w tzw. fazie edycji) miałyby następować programowanie korekty i jej artykulacja (Levelt, 1983). Taki sposób przerywania artykulacji i wprowadzania korekty przedstawia też rysunek 8. Główna zasada przerywania mówienia zakłada priorytetowy charakter poprawności względem płynności mówienia (Seyfeddinipur i in., 2008).

W rozdziale 3. przywołano argumenty na rzecz tego, że zasada ta nie jest uniwersalna i nie ma charakteru obligatoryjnego, gdyż znaleźć można wiele przykładów reperacji, powstających bez natychmiastowej interupcji i przy braku odrębnej fazy edycji (Blackmer i Mitton, 1991). Jak wykazują Seyfeddinipur i in. (2008) ludzie częściej preferują płynność ponad poprawnością, przez co mają tendencję, aby kontynuować artykulację do czasu przygotowania korekty zamiast przerywać ją natychmiastowo po wykryciu błędu. Jednocześnie autorzy ci dopuszczają możliwość istnienia okoliczności, w których osoby tworzące wypowiedź, starają się zmaksymalizować dokładność (kosztem płynności). W takich okolicznościach mówiący dążyliby do jak najszybszego zatrzymania błędnej wypowiedzi oraz wprowadzenia korekty, działając zgodnie z główną zasadą przerywania mówienia. Choć trudno przewidzieć, w jaki sposób warunki takie powstają w naturalnej sytuacji rozmowy i jakie elementy kontekstu (presja czasowa, priorytet przekazywanej informacji, osoba rozmówcy itp.) je tworzą, do pewnego stopnia można uzyskać je w warunkach laboratoryjnych.

Vasić i Wijnen (2005) dowiedli, że przy pomocy manipulacji eksperymentalnej można przekierować skupienie (*focus*) monitora na pewne aspekty wypowiedzi, zmniejszając jednocześnie jego wrażliwość na inne kryteria poprawności. Podobnie Postma i Kolk (1990, za: Postma i Oomen, 2005) wykazali, że uczestnicy badania poinstruowani, aby unikać jakichkolwiek błędów popełniają mniej przejęzyczeń w zadaniu łamańców językowych niż osoby, które otrzymały standardową instrukcję odczytania podanych wyrażen. Co więcej, pod wpływem instrukcji podkreślającej dążenie do poprawności, osoby badane częściej poprawiają popełnione przejęzyczenia niż w warunku standardowej instrukcji. Na tej podstawie można wnioskować, że zorganizowanie odpowiednich warunków eksperymentalnych (w tym instrukcji podkreślającej poprawianie błędów) umożliwia uzyskanie priorytetu poprawności, skłaniającego osoby badane do wkładania wysiłku w korygowanie błędów lub ich unikanie (potencjalnie większy wysiłek może być wkładany w

monitorowanie wewnętrzne i dokonywanie reperacji ukrytych). Idąc dalej, wprowadzenie warunków, w których nie da się uniknąć błędu, może sprzyjać nastawieniu na dokonanie natychmiastowej interupcji i wprowadzenie jak najszybszej na korekty. Z kolei warunki pozwalające na uniknięcie błędu mogą sprzyjać tworzeniu reperacji ukrytych.

Inną kwestią jest to, w jaki sposób warunki te sprzyjają powstawaniu błędów w mówieniu i stymulowaniu procesów kontrolnych odpowiedzialnych za reperacje. Alternatywą dla analizy reperacji dokonywanych na spontanicznych lub eksperymentalnie indukowanych przejrzyczeniach są metody wprost ukierunkowane na symulowanie autokorekcji. Na przykład van Wijk i Kempen (1987) opracowali metodę wywoływania reperacji (*repair elicitation paradigm*) poprzez symulowanie błędów w opisywaniu obrazków: osoba badana opisuje obrazek, który w czasie rzeczywistym ulega przekształceniu, powodując że wygenerowany opis przestaje być adekwatny i konieczne jest jego skorygowanie. Z kolei Boland i in. (2005) opracowali metodę, w której planowana wypowiedź opisująca jeden z obiektów na obrazku, w momencie realizacji planu przestaje być aktualna, gdyż zmianie ulega element kontekstowy, stanowiący punkt odniesienia dla opisu obiektu docelowego<sup>127</sup>. Metody te nie wywołują więc przejrzyczeń, jak np. metoda SLIP (zob. rozdział 3.) albo łamańce językowe, ale wymuszają korygowanie swoich wypowiedzi ze względu na ich nieadekwatność względem zmieniającego się przedmiotu opisu (najczęściej obrazka).

W procedurze zaprojektowanej w niniejszej pracy, osoby badane mają za zadanie powtarzać słyszane zdania zbudowane w szyku SVO<sup>128</sup>. W momencie rozpoczęcia mówienia wyświetlany jest obrazek, który ilustruje treść wypowiedzianego zdania. Symulacja pomyłek leksykalnych oraz indukowanie reperacji odbywa się poprzez wyświetlanie obrazków, które prezentują wykonawców czynności (podmiot zdania) lub przedmiot czynności (dopełnienie) niezgodne z treścią powtarzanego zdania. W ten sposób wprowadzone zostają dwa warunki (1) warunek błędnego podmiotu oraz (2) warunek błędnego dopełnienia. Reperacje w obu warunkach nie wymagają zmiany szyku, polegając jedynie na zamianie *lemmy* (por. van Wijk i Kempen, 1987) na początku zdania (podmiot) lub na jego końcu (dopełnienie).

1. Warunek błędnego podmiotu. Błąd uzewnętrznia się już na samym początku realizacji planu, gdyż wykonawca czynności na obrazku jest inny niż w usłyszonym (i powtarzonym) zdaniu. W ten sposób uczestnik badania nie ma szans na poprawienie

127 Na przykład kiedy opisywany jest *szary kwadrat* będący w sąsiedztwie *białego kwadratu*, nie potrzeba dodawać innych określeń, aby odbiorca komunikatu zrozumiał, który obiekt stanowi przedmiot wypowiedzi. Kiedy jednak *biały kwadrat* zostanie nagle zastąpiony *ciemnoszarym kwadratem*, konieczne jest dodanie informacji o poziomie jasności kwadratu docelowego (np. *jasny, szary kwadrat*), aby wyraźnie odróżnić opisywany obiekt od kontekstu, który uległ zmianie.

128 Na początku zdania występuje rzeczownik w funkcji podmiotu, kolejny wyraz to czasownik pełniący funkcję orzeczenia, a na końcu zdania znów znajduje się rzeczownik lub wyrażenie przyimkowe pełniące funkcję dopełnienia.

błędu w planie przedartykulacyjnym (reperacja ukryta) i konieczne jest wprowadzenie korekty zewnętrznej poprzez restart całego zdania. Jednocześnie pozostała część planu wypowiedzi jest poprawna, co oznacza że w przypadku zatrzymania wypowiedzi po wykryciu błędu, konieczna jest tylko zamiana *lemmy* rozpoczynającej zdanie, bez modyfikacji pozostałych elementów przechowywanych w buforze artykulacyjnym. Zgodnie z modyfikacją modelu pętli percepcyjnej, zaproponowaną przez Hartsuikera i Kolka (2001; por. Levelt, 1989), procesy inicjujące interrupcję reparandum oraz zaplanowanie korekty nie są sekwencyjne, lecz równoległe. Inaczej mówiąc, w wyniku wykrycia błędu przez monitor, kieruje on dwutorowym działaniem zatrzymania artykulacji pierwotnego planu, który zawiera błąd (już wypowiedziany lub jeszcze znajdujący się w buforze artykulacyjnym) oraz zaplanowania korekty i jej realizacji. Tym samym w warunku błędnego podmiotu nie ma potrzeby, aby opóźnić interrupcję, gdyż w momencie wykrycia niezgodności między osobą na obrazku a agensem czynności w powtarzonym zdaniu, aktywowane zostaje pojęcie leksykalne (rozpoznanie postaci na obrazku), które może uruchomić proces selekcji leksykalnej, przebiegający równoległe z zatrzymywaniem bieżącej artykulacji. W ten sposób, reperacja może być gotowa już w momencie interrupcji lub tuż po niej. Taki warunek powinien wzmacniać preferencję natychmiastowej interrupcji i skłaniać osoby badane do jak najwcześniejszego przerywania mówienia. Tym samym za pomocą tej procedury możliwe będzie zbadanie, jak szybko uczestnicy badania dokonują interrupcji i czy istnieją w tym zakresie między nimi różnice wynikające z wieku i/lub poziomu posiadanych funkcji wykonawczych.

2. Warunek błędnego dopełnienia. W momencie rozpoczęcia artykulacji przygotowanego planu nie ujawnia się jeszcze błąd, gdyż agens czynności z usłyszanego zdania oraz postać na obrazku są ze sobą zbieżne (pierwszy oraz drugi wyraz są poprawne względem treści obrazka). Niezgodność dotyczy natomiast obiektu czynności na obrazku oraz dopełnienia artykułowanego na ostatniej pozycji zdania. Według Hartsuikera i Kolka (2001) końcowe części przechowywanej w buforze artykulacyjnym wypowiedzi mają większą szansę na wykrycie w nich błędu jeszcze przez monitorowanie wewnętrzne oraz reperację jeszcze przed uzewnętrznieniem się tego błędu. Dzieje się tak, gdyż w przypadku końcowych części buforowanej frazy monitorowanie wewnętrzne ma do dyspozycji więcej czasu na analizę, dzięki czemu może zainicjować przeformułowanie tej części planu fonetycznego, zanim zostanie on zrealizowany (por. Boland i in., 2005). Zgodnie z modelem Hartsuikera i Kolka (2001)

w chwili rozpoznania niezgodności w warunku błędnego dopełnienia możliwe powinno być uruchomienie procesu selekcji leksykalnej oraz dokonanie korekty na poziomie buforowanego planu przedartykulacyjnego jeszcze w trakcie realizacji jego poprawnej części (podmiot i orzeczenie zdania) w taki sposób, aby błędny wyraz (dopełnienie) w ogóle nie został wypowiedziany. Tym samym warunek błędnego dopełnienia – w przeciwieństwie do warunku błędnego podmiotu – stwarza szansę na dokonanie reperacji ukrytych, które według klasyfikacji Postmy i Kolka (1993) mogą przybierać jedną z dwóch form:

- a. odtworzenie (*restart*) – zatrzymanie artykulacji przed wypowiedzeniem błędnego wyrazu i wznowienie mówienia od momentu poprzedzającego interupcję (od początku zdania lub od orzeczenia). Reperacja taka odtwarza więc w całości lub częściowo plan zrealizowany do momentu interupcji oraz uzupełnia go o poprawną wersję skorygowanego wyrazu.
- b. odroczenie (*postponement*) – spowolnienie lub zatrzymanie artykulacji przed wypowiedzeniem błędnego wyrazu oraz kontynuowanie wypowiedzi (lub jej wznowienie w razie przerwania) bez jakiegokolwiek odtwarzania lub powtarzania wcześniej zrealizowanych części planu wypowiedzi. Wypowiadane zdanie nie ujawnia w warstwie tekstowej żadnego zaburzenia (wypowiedź jest w pełni zgodna z prezentowanym obrazkiem), ale na poziomie artykulacji zauważyć można spowolnienie tempa, wystąpienie dłuższej pauzy pomiędzy wyrazami lub w obrębie artykulacji któregoś z wyrazów poprzedzających błędny wyraz.

Poza wskazanymi powyżej reperacjami ukrytymi możliwe jest również wypowiedzenie całego zdania wraz z błędnym wyrazem na końcu (bez interupcji) i poprawienie go już w postaci jawnej korekty.

Zastosowane zadanie symulujące dokonywanie reperacji, zostało szczegółowo przedstawione w sekcji 5.4.2.1.3. tego rozdziału. Przy jego użyciu – dzięki kontroli miejsca występowania błędu w strukturze zdania – możliwe będzie przyjrzenie się sprawności w dokonywaniu reperacji zarówno pod względem szybkości interupcji oraz umiejętności dokonywania reperacji ukrytych. Jego wyniki posłużą w sformułowaniu odpowiedzi na pytanie dotyczące tego, czy wraz z wiekiem zmienia się sposób wprowadzania korekty do błędnej wypowiedzi oraz w jakim stopniu różnice w dokonywaniu korekt przez ludzi w różnym wieku są związane z poziomem kontroli wykonawczej. Szczegółowe pytania badawcze i hipotezy związane z tym problemem przedstawione zostały w podrozdziale 5.2.

### **5.1.4 Podsumowanie**

Reasumując, przejęzyczenie jest podstawowym (modelowym) przejawem zakłóceń i kontroli procesu mówienia, wykorzystywanym w badaniach empirycznych. Podczas przejęzyczenia i jego korekty ujawnia swoje działanie proces monitorowania, który odpowiada za detekcję błędu i zainicjowanie reperacji. Z tego względu badacze produkcji mowy analizują ilościowe i jakościowe aspekty spontanicznych oraz eksperymentalnie indukowanych przejęzyczeń. Jednocześnie elementarne procesy występujące w przejęzyczeniach i ich reperacjach można z powodzeniem symulować w kontrolowanych zadaniach eksperymentalnych, aby uwypuklić działanie mechanizmów kontroli na poszczególnych etapach programowania, realizacji oraz rewizji i reperacji wypowiedzi. Metodologia ta – oparta na zadaniach odnoszących się do elementarnych procesów selekcji leksykalnej, monitorowania oraz dokonywania reperacji – została wykorzystana w niniejszej pracy (szczegóły w części 5.4.2.1.), aby odpowiedzieć na pytania badawcze zaprezentowane poniżej.

## **5.2 Problemy badawcze i hipotezy**

Jak wykazano w poprzedniej części rozprawy, poznawcze modele produkcji mowy naczelną rolę w kontroli procesu mówienia przypisują monitorowi (lub monitorom) odpowiedzialnemu za sprawdzanie (monitorowanie) poprawności generowanej wypowiedzi oraz jej zgodności z intencją komunikacyjną nadawcy. Podstawowym, modelowym dla badań psycholingwistycznych przejawem poznawczej kontroli procesu mówienia jest przejęzyczenie i jego korekta, w których manifestują się behawioralne wskaźniki mechanizmów kontrolnych produkcji mowy. Dlatego w badaniach z tego obszaru analizuje się spontaniczne lub eksperymentalnie indukowane przejęzyczenia oraz ich reperacje, a także towarzyszące im niepełności w postaci pauz, powtórzeń, falstartów itp.

Mimo iż, propozycje teoretyczne na temat tego, w jaki sposób przebiega monitorowanie wewnętrzne różnią się między sobą (co ma podłoże w odmiennym podejściu do organizacji procesu produkcji mowy, zob. rozdział 2 i 3), to – niezależnie od orientacji teoretycznej<sup>129</sup> – coraz częściej w badaniach podnosi się związki monitorowania procesu mówienia z kontrolą wykonawczą. Na przykład Shao i in. (2012) wykazali, że czas reakcji werbalnej w nazywaniu czynności (prezentowanych na obrazkach) istotnie negatywnie koreluje z miarą odświeżania pamięci roboczej, a czas reakcji werbalnej w nazywaniu przedmiotów i czynności istotnie negatywnie koreluje z miarą hamowania (czas reakcji w

---

<sup>129</sup> Zob. dyskusja na łamach *Journal of Cognition* 3(1) 2020 toczona przez Roelofsa i Nozari (Nozari; 2020; Roelofs, 2020a; 2020b; por. Lind i Hartsuiker, 2020).

zadaniu sygnału stop rozwiązywanym manualnie). Wardlow (2013) wykazała z kolei, że prezentowany przez uczestników badania poziom funkcji wykonawczych koreluje z adekwatnością podawanych przez nich informacji w zadaniu nazywania. Przykładowo, im wyższe wyniki badani uzyskali w zakresie kontroli hamowania, tym rzadziej używali nieistotnych – z punktu widzenia rozmówcy – informacji na temat opisywanych obiektów.

Jednocześnie, wciąż niewiele badań ukierunkowanych jest na poznanie rozwojowych aspektów kontroli procesu mówienia. Dotychczasowe doniesienia pokazują jednak dynamiczne zmiany zachodzące w tym zakresie w ontogenezie. Wiadomo na przykład, że szybkość nazywania obrazków rośnie wraz z wiekiem w okresie późnego dzieciństwa, osiągając najwyższy wskaźnik w okresie dorosłości (Rosinski, 1977). Wraz z wiekiem spada też liczba popełnianych przez dzieci błędów semantycznych, a rośnie odsetek korekt tych błędów (Hanley i in., 2016). Ponadto monitorowanie wewnętrzne wyrazów zwiększa swoją sprawność w okresie późnego dzieciństwa (Sasisekaran i Weber-Fox, 2012).

Nadal jednak można mówić także o niedostatku badań podejmujących problem różnic indywidualnych w kontroli procesu mówienia, jak i zmian rozwojowych jakim podlega ten proces i mechanizmy jego kontroli. Co więcej, brakuje badań które łączyłyby obie perspektywy weryfikując związki funkcji wykonawczych z kontrolą procesu mówienia na tle postępującego wraz z wiekiem progresu w zakresie kontroli wykonawczej i sprawności w posługiwaniu się mową. Z tego względu postanowiono podjąć ten problem w niniejszym projekcie, formułując następujące, podstawowe pytania badawcze: **Czy kontrola procesu mówienia zmienia się wraz z wiekiem od okresu późnego dzieciństwa do dorosłości? Jeśli tak, to (1) w jaki sposób osoby w odmiennych grupach wiekowych różnią się pod względem wybranych aspektów kontroli procesu mówienia oraz (2) w jakim stopniu sprawność wybranych aspektów kontroli procesu mówienia wiąże się z poziomem funkcji wykonawczych u tych osób?**

### **5.2.1 Pytania badawcze**

W związku z tym, że przyjętym w niniejszym projekcie modelowym przejawem kontroli procesu mówienia jest przejęzyczenie i jego reperacja, obejmujące bardziej elementarne zjawiska, które można względnie niezależnie od siebie symulować, postawiono kolejne, bardziej szczegółowe pytania badawcze dotyczące różnic między grupami wiekowymi w kontroli procesu mówienia wynikające z wieku i manipulacji eksperymentalnej (pytania od 1.1.1 do 1.3.2) oraz związków kontroli procesu mówienia z poziomem funkcji wykonawczych (pytania od 2.1.1. do 2.3.2).

1. Grupa pytań o różnice w kontroli procesu mówienia związane z wiekiem oraz zastosowaną manipulacją eksperymentalną:

Pytanie 1.1. Czy wraz z wiekiem, u osób w okresie od późnego dzieciństwa przez adolescencję do wczesnej dorosłości, zmienia się poziom kontroli selekcji leksykalnej?

Pytanie 1.1.1. Czy skuteczność selekcji leksykalnej zmienia się wraz z wiekiem<sup>130</sup>?

Pytanie 1.1.2. Czy skuteczność selekcji leksykalnej spada pod wpływem zewnętrznej dystrakcji semantycznej u osób w okresie późnego dzieciństwa, adolescencji i dorosłości?

Pytanie 1.1.3. Czy wraz z wiekiem zmienia się poziom odporności selekcji leksykalnej na zewnętrzną dystrakcję semantyczną?

Pytanie 1.1.4. Czy szybkość<sup>131</sup> selekcji leksykalnej zmienia się wraz z wiekiem?

Pytanie 1.1.5. Czy szybkość selekcji leksykalnej u osób w okresie późnego dzieciństwa, adolescencji i dorosłości zmienia się pod wpływem zewnętrznej dystrakcji semantycznej?

Pytanie 1.1.6. Czy wielkość efektu interferencji semantycznej zmienia się wraz z wiekiem?

Pytanie 1.2. Czy wraz z wiekiem, w okresie od późnego dzieciństwa przez adolescencję do wczesnej dorosłości, zmienia się efektywność automonitorowania procesu mówienia?

Pytanie 1.2.1. Czy dokładność<sup>132</sup> monitorowania wewnętrznego zmienia się wraz z wiekiem?

Pytanie 1.2.2. Czy szybkość monitorowania wewnętrznego zmienia się wraz z wiekiem?

Pytanie 1.2.3. Czy szybkość monitorowania wewnętrznego u osób w wyróżnionych grupach wiekowych uzależniona jest od pozycji wykrywanego elementu w strukturze monitorowanego wyrazu?<sup>133</sup>

Pytanie 1.2.4. Czy szybkość monitorowania wewnętrznego u osób w wyróżnionych grupach wiekowych zależy od długości monitorowanego wyrazu?

---

130 Wszystkie pytania i hipotezy badawcze dotyczące wieku odnoszą się do okresu od późnego dzieciństwa do wczesnej dorosłości (przedział wiekowy badanych osób to 7;8-25;2, zob. 5.4.1.3.).

131 Szybkość selekcji leksykalnej rozumiana tu jest jako ilość czasu potrzebnej na skuteczne (poprawne) dokonanie wyboru właściwej jednostki leksykalnej (zob. tabela 3).

132 Dokładność (*accuracy*) rozumiana jest tu jako precyzja i trafność działania monitora, wyrażająca się w procencie poprawnych odpowiedzi w zadaniu *Monitorowanie sylab* (por. tabela 3).

133 Pytanie o zależność między szybkością monitorowania, a pozycją sylaby w wyrazie dotyczy tego, jak dużo czasu „monitor” potrzebuje na dokonanie inspekcji wewnętrznego planu produkcji mowy, aby wykryć sylaby na różnych pozycjach (nagłos, śródgłos, wygłos). Nie jest to pytanie o różne tempo monitorowania w zależności od miejsca sylaby w strukturze, ale o zmiany w czasie efektywnego monitorowania wynikające z długości tekstu do sprawdzenia.

Pytanie 1.3. Czy wraz z wiekiem, w okresie od późnego dzieciństwa przez adolescencję do wczesnej dorosłości, zmienia się sposób korygowania błędnych wypowiedzi?

Pytanie 1.3.1. Czy wraz z wiekiem zmienia się szybkość dokonywania interupcji w błędnej wypowiedzi?

Pytanie 1.3.2. Czy wraz z wiekiem zmienia się sposób dokonywania reperacji?

2. Grupa pytań o związki między kontrolą procesu mówienia a poziomem funkcji wykonawczych:

Pytanie 2.1. W jakim stopniu poziom kontroli selekcji leksykalnej wiąże się z poziomem funkcji wykonawczych u osób w okresie późnego dzieciństwa, adolescencji i dorosłości?

Pytanie 2.1.1. Czy zachodzą związki między odpornością selekcji leksykalnej na dystrakcję semantyczną a poziomem funkcji wykonawczych u osób w wybranych grupach wiekowych? Jeśli tak, to jakie?

Pytanie 2.1.2. Czy zachodzą związki między wielkością efektu interferencji semantycznej a poziomem funkcji wykonawczych u osób w wybranych grupach wiekowych? Jeśli tak, to jakie?

Pytanie 2.2. W jakim stopniu automonitorowanie procesu mówienia wiąże się z poziomem funkcji wykonawczych u osób w okresie późnego dzieciństwa, adolescencji i dorosłości?

Pytanie 2.2.1. Czy zachodzą związki między dokładnością monitorowania wewnętrznego a poziomem funkcji wykonawczych u osób w wybranych grupach wiekowych? Jeśli tak, to jakie?

Pytanie 2.2.2. Czy zachodzą związki między szybkością monitorowania wewnętrznego a poziomem funkcji wykonawczych u osób w wybranych grupach wiekowych? Jeśli tak, to jakie?

Pytanie 2.3. W jakim stopniu sposób dokonywania reperacji wiąże się z poziomem funkcji wykonawczych u osób w okresie późnego dzieciństwa, adolescencji i dorosłości?

Pytanie 2.3.1. Czy zachodzą związki między szybkością dokonywania interupcji błędnych wypowiedzi a poziomem funkcji wykonawczych u osób w wybranych grupach wiekowych? Jeśli tak, to jakie?

Pytanie 2.3.2. Czy zachodzą związki między częstością stosowania reperacji wewnętrznych przez strategię odraczania, a poziomem funkcji wykonawczych u osób w wybranych grupach wiekowych? Jeśli tak, to jakie?

Uwzględniając wciąż niewielką liczbę prac badawczych w psycholingwistyce łączących perspektywę rozwojową z badaniem różnic indywidualnych w kontroli procesu mówienia,

prezentowane w niniejszej rozprawie rozważania na ten temat mają szansę rozszerzyć dotychczasowy stan wiedzy.

## **5.2.2 Hipotezy**

### **5.2.2.1 Grupa hipotez dotyczących pytań badawczych 1.1.1-1.3.2**

Hipotezy z tej grupy dotyczą różnic w kontroli procesu mówienia związanych z wiekiem oraz zastosowaną manipulacją eksperymentalną i odpowiadają na pytania badawcze 1.1.1-1.3.2

#### **Hipoteza 1.1. Poziom kontroli selekcji leksykalnej wzrasta wraz z wiekiem w okresie od późnego dzieciństwa przez adolescencję do wczesniej dorosłości.**

Hipoteza 1.1.1. Skuteczność selekcji leksykalnej wzrasta wraz z wiekiem.

Hipoteza 1.1.2. Pod wpływem zewnętrznej dystrakcji semantycznej skuteczność selekcji leksykalnej obniża się u przedstawicieli wszystkich badanych grup wiekowych.

Hipoteza 1.1.3. Poziom odporności selekcji leksykalnej na zewnętrzną dystrakcję semantyczną wzrasta wraz z wiekiem.

Hipoteza 1.1.4. Szybkość selekcji leksykalnej wzrasta wraz z wiekiem.

Hipoteza 1.1.5. Szybkość selekcji leksykalnej zmniejsza się pod wpływem zewnętrznej dystrakcji semantycznej u osób w różnym wieku.

Hipoteza 1.1.6. Wielkość efektu interferencji semantycznej maleje wraz z wiekiem.

#### **Uzasadnienie:**

Selekcja leksykalna to proces wyboru prawidłowego (tj. właściwego pod względem referencji) wyrazu ze słownika umysłowego. W modelu Levelta termin ten odnosi się do poziomu produkcji mowy, na którym dokonuje się wybór lemmy odpowiadającej aktywowanemu pojęciu leksykalnemu (Levelt, 1998; 1999a; 1999b; Levelt i in., 1999;). Rozpoczyna ona dwustopniowy proces uzyskiwania dostępu leksykalnego, w którym reprezentacja semantyczna (pojęcie leksykalne) otrzymuje swoją formę językową (wyraz fonologiczny). W przypadku generowania wypowiedzi dłuższych niż pojedynczy wyraz, w których zachodzą relacje syntagmatyczne pomiędzy elementami wypowiedzianego tekstu, selekcja leksykalna stanowi część kodowania gramatycznego, które organizuje wypowiedziane wyrazy według relacji składniowych.

Selekcja leksykalna i kodowanie gramatyczne mają kluczowe znaczenie dla realizacji intencji komunikacyjnej, gdyż na tym poziomie może dojść do błędów substytucji całych wyrazów w planie paradygmatycznym (np. wypowiedzenie „prawo” zamiast „lewo” itp.) lub syntagmatycznym (zamiana miejscami w zdaniu wyrazów należących do tych samych kategorii gramatycznych, np. zamiana podmiotu i dopełnienia „kot gonił psa” skutkująca odwróceniem ról semantycznych realizowanych przez te wyrazy) (Garrett, 1976; 1980, Levelt, 1989; Bręński, 2013).

Kontrola wyboru odpowiednich jednostek leksykalnych jest więc istotnym elementem w kontroli całego procesu mówienia (Vigliocco i in., 1997; Levelt, 1999a). Według Levelta i in. (1999; Roelofs, 2020a) za jej powodzenie odpowiada nie tylko aktywacja właściwych jednostek leksykalnych, ale także ich weryfikacja przez mechanizmy kontrolne działające na poziomie leksykonu. Stopień w jakim selekcja leksykalna prowadzi do realizacji (wyboru a następnie artykulacji) właściwego wyrazu będzie dalej określana jako jej skuteczność.

Na podstawie badania Hanleya i in. (2016) można wnioskować, że skuteczność selekcji leksykalnej wzrasta na przełomie średniego i późnego dzieciństwa, co objawia się zmniejszeniem błędów w nazywaniu obrazków. W okresie późnego dzieciństwa i adolescencji wzrasta także szybkość nazywania obrazków (Rosinski, 1977), co można – przynajmniej częściowo – przypisać większej szybkości selekcji leksykalnej. Czas efektywnej selekcji leksykalnej ma przełożenie na płynność mówienia, gdyż trudność w selekcji właściwych jednostek leksykalnych prowadzi do wystąpienia przerw w mówieniu w postaci pauz wypełnionych (Hartsuiker i Notebaert, 2010), a nawet zjawiska „na końcu języka” (Vigliocco i in., 1997; Levelt i in. 1999) Sukcesywne zwiększanie płynności mowy (w tym wzrost tempa mówienia) w okresie dzieciństwa (Sabin i in., 1979) także można wiązać z większą efektywnością pod względem temporalnego aspektu selekcji leksykalnej.

Selekcja leksykalna jest wrażliwa na kontekst semantyczny. W korpusach błędów mowy znaleźć można przykłady przejęzyczeń w postaci substytucji wyrazów wynikające z tego, że ktoś w trakcie mówienia równolegle przetwarza inną, niezwiązaną z wypowiedzią treść (myśli na inny temat, słyszenie rozmowy w tle lub przeczytanie jakiegoś napisu) (Harley, 1984; 2001). Eksperymentalna manipulacja kontekstem mówienia poprzez wprowadzenie dystrakcji semantycznej do zadania nazywania obrazków powoduje spowolnienie czasu reakcji (Levelt, 1993; Schriefers i in., 1990). Opóźnienie w nazywaniu interpretuje się zazwyczaj jako wydłużenie czasu selekcji leksykalnej, związane z kosztem przewycięzania konkurencji między reprezentacjami wzbudzonymi równocześnie przez obiekt na obrazku (wyraz docelowy) i dystraktor (wyraz zakłócający). Poza wydłużeniem

czasu efektywnej selekcji leksykalnej<sup>134</sup> możliwy jest też spadek jej skuteczności (Jescheniak i in., 2020), jednak w przypadku zdrowych dorosłych uczestników badań jest on na tyle marginalny, że w większości publikacji raportuje się go w statystykach opisowych najczęściej jako odsetek błędnych odpowiedzi (*error rate*) i nie poddaje dalszym analizom statystycznym. Ten wysoki poziom poprawności nazywania w warunkach interferencji semantycznej pokazuje, że selekcja leksykalna u osób dorosłych jest wysoce odporna na dystrakcję semantyczną. Inaczej jednak jest w przypadku grup klinicznych. Jak wykazali Harvey i in. (2019) manipulacja kontekstem semantycznym i nasilenie efektu interferencji<sup>135</sup> może zwiększać odsetek błędów leksykalnych w odpowiedziach osób z afazją.

W przypadku dzieci (okres szkolny) efekt interferencji (opóźnienie nazywania w wyniku dystrakcji) może być nawet większy niż u osób dorosłych (Rosinski, 1977), co – w połączeniu z ogólnie wolniejszym tempem nazywania obrazków oraz mniejszą skutecznością selekcji leksykalnej (Hanley i in., 2016) – może skutkować istotnym obniżeniem poprawności nazywania pod wpływem dystrakcji semantycznej. Można więc spodziewać się, że odporność na dystrakcję semantyczną będzie wraz z wiekiem (na przestrzeni późnego dzieciństwa i adolescencji) wykazywać tendencję wzrostową, osiągając najwyższy poziom w okresie dorosłości.

Zastosowanie tej samej procedury badawczej (zadanie nazywania obrazków z manipulacją zewnętrzną dystrakcją semantyczną) w trzech grupach wiekowych (późne dzieciństwo, adolescencja, wczesna dorosłość) pozwoli ustalić zakres zmian wynikających z wieku pod względem skuteczności selekcji leksykalnej, jej efektywności czasowej oraz odporności na dystrakcję. Przewidywane rezultaty badania zostały przedstawione w postaci hipotez 1.1.1.-1.1.6., sformułowanych na podstawie dotychczasowych badań oraz przyjętego modelu teoretycznego. Oczekuje się przede wszystkim zreplicowania<sup>136</sup> efektu interferencji semantycznej w wypowiedziach wszystkich grup wiekowych. Spodziewane jest także przełożenie się tego efektu na poprawność wykonania zadania w postaci różnic w odsetku

134 Dla oznaczenia, że chodzi o selekcję leksykalną skutkującą poprawnym nazywaniem obrazków, będzie stosowane określenie „czas efektywnej selekcji leksykalnej”.

135 W badaniu nie prezentowano osobom badanym bodźców dystrakcyjnych, lecz nazywane obrazki wyświetlane były w losowej kolejności, co stwarzało okazję do powstawania serii obrazków należących do tej samej kategorii semantycznej. Nazywanie obrazków powiązanych semantycznie jest – oprócz bezpośredniej dystrakcji – jednym ze sposobów wywoływania interferencji semantycznej (Harvey, i in., 2019).

136 Wedle posiadanej przez autora wiedzy badania przy wykorzystaniu procedury nazywania obrazków z dystrakcją nie były do tej pory prowadzone w języku polskim, ani z osobami dla których jest on językiem ojczystym. Wyjątkiem jest badanie Kujalowicz i Zajdler (2009), w którym osoby posługujące się językiem polskim jako ojczystym (L1) nazywały obrazki w języku chińskim (L3) otrzymując dystrakcję w języku angielskim (L2) lub polskim (L1). W badaniu tym produkcja mowy prowadzona była jednak w języku obcym (L3), a ojczysty język polski służył jedynie do wprowadzenia dystrakcji w jednym z wariantów zadania. Efekt interferencji zauważalny był w produkcji mowy w języku chińskim zarówno, kiedy dystrakcja podawana była w języku angielskim, jak i w języku polskim.

poprawnie nazwanych obrazków w dwóch różnych warunkach (nazywanie vs. nazywanie z dystrakcją) w odmiennych grupach wiekowych.

**Hipoteza 1.2. Efektywność automonitorowania procesu mówienia wzrasta wraz z wiekiem w okresie od późnego dzieciństwa przez adolescencję do wczesniej dorosłości.**

Hipoteza 1.2.1. Monitorowanie wewnętrzne zwiększa swoją dokładność wraz z wiekiem.

Hipoteza 1.2.2. Szybkość monitorowania wewnętrznego wzrasta wraz z wiekiem.

Hipoteza 1.2.3. Szybkość monitorowania wewnętrznego maleje wraz z zajmowaniem przez wykrywany element kolejnych pozycji w strukturze wyrazu (wydłuża się czas efektywnego monitorowania).

Hipoteza 1.2.4. Szybkość monitorowania wewnętrznego maleje wraz ze wzrostem długości monitorowanego wyrazu (wydłuża się czas efektywnego monitorowania).

**Uzasadnienie:** Automonitorowanie jest procesem kontrolnym produkcji mowy, który odpowiada za weryfikację własnej wypowiedzi nadawcy pod względem jej poprawności oraz zgodności z intencją. Działanie automonitorowania umożliwia wykrywanie i poprawianie błędów (monitorowanie zewnętrzne i wewnętrzne) oraz zapobieganie ich występowaniu (monitorowanie wewnętrzne) (Hartsuiker i Kolk, 2001; Hartsuiker i in. 2005). Monitorowanie zewnętrzne badać można analizując spontaniczne lub eksperymentalnie indukowane przejęzyczenia oraz ich korekty, natomiast o monitorowaniu wewnętrznym często wnioskuje się pośrednio, np. na podstawie ukrytych reperacji (Levelt, 1983; 1989; Postma i Kolk, 1993). W związku z tym, że automonitorowanie nie ogranicza się jedynie do reagowania na błędy, ale działa jako bardziej generalny mechanizm sprawdzający poprawność produkowanego tekstu (por. Abel i in., 2009) jego charakterystykę można badać także w zadaniach niewymagających generowania pomyłek i przejęzyczeń (Wheeldon i Levelt, 1995), uzyskując zbliżone rezultaty do zadań angażujących osoby badane do popełniania pomyłek (Nooteboom i Quené, 2019). Podstawowe wnioski płynące z tych badań dotyczą tego, że monitorowanie wewnętrzne ma dostęp do fonologicznej reprezentacji generowanych wyrazów oraz, że jego działanie ma charakter przyrostowy, tj. sprawdza generowany tekst sukcesywnie weryfikując najpierw segmenty występujące na wcześniejszych pozycjach, następnie przechodząc do pozycji dalszych (Schiller, 2005; Wheeldon i Levelt, 1995). Uwidacznia się to w czasie potrzebnym do wykrycia elementów subleksykalnych (fonemów, sylab) w generowanych

wyrazach: elementy znajdujące się na dalszych pozycjach wykrywane są później w porównaniu do elementów znajdujących się na początku wyrazu.

Podstawowym oczekiwanym rezultatem projektowanego badania jest replikacja efektu pozycji wyszukiwanego elementu (sylaby) w wyrazie oraz wykazanie, że wyrazy krótsze monitorowane są w proporcjonalnie krótszym czasie, co wynika z założeń przyjętego modelu teoretycznego (Levelt, 1999b; Levelt i in. 1999; Roelofs, 2005a). Jednocześnie przewiduje się, że wraz z wiekiem będzie wzrastała efektywność monitorowania wewnętrznego czyli (1) wzrośnie dokładność oraz (2) skróceniu ulegał będzie czas efektywnego monitorowania wewnętrznego (większa szybkość monitorowania), jako przejaw zmian rozwojowych obejmujących ten proces. Wniosek ten nasuwa się na podstawie ogólnego twierdzenia, że dzieci u progu wieku szkolnego wchodzą w okres osiągania pełnej kompetencji językowej, który rozciąga się do połowy, a nawet końca późnego dzieciństwa (Kurcz, 2005) oraz na podstawie wyników badania Sasisekaran i Weber-Fox (2012), które wykazały, że proces monitorowania u dzieci w wieku 7-13 lat zwiększa swoją dokładność. Również dane z analizy spontanicznych wypowiedzi rejestrowanych w naturalnych warunkach codziennych interakcji potwierdzają zasadność oczekiwania, że zdolność monitorowania rozwija się wraz z wiekiem w okresie dzieciństwa (Evans, 1985; Jaeger, 2005). Trudno jednak oddzielić na ile jest to progres w zakresie ogólnej efektywności monitora mowy, a na ile usprawnieniu ulega monitorowanie zewnętrzne lub wewnętrzne. To pierwsze manifestuje się już od najwcześniejszych etapów posługiwania się mową w 2. roku życia dziecka (Clark, 1978; Laakso, 2010), choć jednocześnie przed 4. rokiem życia obserwuje się jeszcze niestabilność działania procesów monitorujących i korygujących wyartykułowany tekst (MacDonald i in., 2012; por. Jaeger, 2005). O rozwoju tego drugiego niewiele można powiedzieć na podstawie dostępnej literatury przedmiotu, jednak ustalenia dotyczące rozwoju mowy wewnętrznej<sup>137</sup> (Geva i Fernyhough, 2019) pozwalają domniemywać, że monitorowanie wewnętrzne wykształca się we wczesnym i średnim dzieciństwie, po czym doskonalą się w okresie późnego dzieciństwa i adolescencji.

Zastosowanie tej samej procedury badawczej (zadanie podwójne: bezgłośnie nazywanie obrazków z jednoczesnym wykrywaniem sylab w strukturze wyrazów nazywających obrazki) w trzech grupach wiekowych (późne dzieciństwo, adolescencja, wczesna dorosłość) pozwoli ustalić zakres różnic w efektywności monitorowania wewnętrznego wynikających z wieku.

---

<sup>137</sup> Fernyhough (2004; Geva i Fernyhough, 2019) nawiązuje do konceptualizacji mowy wewnętrznej (*inner speech*) według teorii Wygotskiego (1934/1989), jednocześnie rozwijając ją i próbując odnieść do psycholingwistycznych badań nad monitorowaniem produkcji mowy.

### **Hipoteza 1.3. Sposób korygowania błędnych wypowiedzi ulega zmianom wraz z wiekiem w okresie od późnego dzieciństwa przez adolescencję do wcześniej dorosłości.**

Hipoteza 1.3.1. Szybkość dokonywania interupcji w błędnej wypowiedzi wzrasta wraz z wiekiem.

Hipoteza 1.3.2. Częstość stosowania reperacji ukrytych (strategia odraczania) wzrasta wraz z wiekiem.

#### **Uzasadnienie:**

Cytowane do tej pory badania dostarczają danych na temat wzrostu kontroli procesu mówienia wraz z wiekiem poprzez spadek liczby popełnianych błędów oraz wzrost odsetka korekt (np. Hanley i in., 2016; Jaegger, 2005). Jest to wspólny efekt doskonalenia się procesów produkcji mowy oraz monitorowania, które wraz z wiekiem coraz szybciej i skuteczniej weryfikuje tworzone wypowiedzi. Choć analiza reperacji jest nieodzownym elementem badania kontroli procesu mówienia do tej pory niewiele uwagi poświęcono różnicom indywidualnym oraz zmienności rozwojowej w samym sposobie dokonywania autokorekt przez typowych użytkowników języka. W związku z tym w niniejszej pracy postanowiono sprawdzić dwa aspekty dokonywania reperacji u osób z różnych grup wiekowych: (1) szybkość dokonywania interupcji w reperacjach zewnętrznych (H.1.3.1.) oraz (2) częstość stosowania reperacji ukrytych dokonywanych według strategii odraczania (H.1.3.2.).

Reperacje przejęzyczeń obecne są w mowie dzieci od momentu, kiedy ich komunikacja przyjmuje rzeczywiście charakter językowy (Hanley i in., 2016; por. Clark, 1978; Laakso, 2010). Zatem bardzo wcześnie pojawia się zdolność zatrzymania artykulacji (kluczowa w przejściu z gaworzenia do mówienia) oraz wprowadzenia korekty. Od tej pory możliwe staje się obserwacyjne badanie, jak wiele reperacji pojawia się w mowie dziecka, jakich błędów dotyczą (które błędy są poprawiane częściej od innych) oraz jaki mają charakter. Jeśli potraktuje się reperację i poprzedzającą ją interupcję jako podstawową reakcję na wystąpienie błędu w mowie (lub jej planie przedartykulacyjnym), można zastanowić się, czy szybkość z jaką podejmowana jest ta reakcja zmienia się wraz z wiekiem. Inaczej mówiąc, rozważane są tu rozwojowe zmiany sprawności procesów kontrolnych polegające na zyskiwaniu biegłości w coraz szybszym reagowaniu na popełniony lub potencjalny błąd w postaci: (1) jak najszybszego zatrzymywania artykulacji po popełnieniu

błędu lub jeszcze w trakcie jego artykulacji (H.1.3.1.); (2) stosunkowo częstego dokonywania reperacji ukrytych (H.1.3.2.).

Analiza spontanicznych konwersacji prowadzonych przez dzieci i dorosłych pokazuje, że bardzo wcześnie ważne staje się to, aby reperacja wykonana była samodzielnie i to jeszcze w tej samej turze konwersacyjnej (Forrester, 2008; Laakso, 2010). Realizując w mowie swoje intencje komunikacyjne dzieci śledzą swoje wypowiedzi oraz ich rezultaty (zachowanie interlokutora) i reagują na powstające niezgodności. Zbyt długie zwlekanie z podjęciem reperacji mogłoby skutkować niezrozumieniem intencji komunikacyjnej przez odbiorcę lub podjęciem przez niego próby skorygowania wypowiedzi nadawcy. Można więc przypuszczać, że już we wczesnym i średnim dzieciństwie następuje progres w reagowaniu na błędy i nieścisłości nie tylko pod względem wzrostu częstości reperacji (Evans, 1985; Jaeger, 2005), ale też szybkości ich podejmowania (zatrzymanie mówienia i dokonanie korekty).

W przypadku osób dorosłych zatrzymywanie mówienia w jak najszybszym momencie po wykryciu błędu jest kwestią kontrowersyjną. Z jednej strony główna zasada przerywania mówienia (*MIR*) postuluje niemal automatyczne interupcje w artykulacji po zarejestrowaniu błędu, z drugiej strony raportowane są także liczne przypadki odraczanych reperacji, które nie są poprzedzane nagłą interupcją, lecz płynnie wprowadzane do strumienia mowy (Blackmer i Mitton, 1991). Choć wyniki badania Seyfeddinipur i in. (2008) wspierają hipotezę odroczenia reperacji (*DIR*), pokazując, że osoby mówiące przedkładają płynność ponad dokładność, to jednocześnie nie wykluczają, że w pewnych okolicznościach priorytetem może być poprawność. Takie okoliczności można uzyskać tworząc odpowiednie warunki eksperymentalne. Spodziewanym efektem takiej organizacji zadania badawczego, w którym istotne jest poprawianie błędów, jest uzyskanie priorytetu poprawności ponad płynnością, który pozwoli sprawdzić w jakim stopniu badane osoby różnią się pod względem szybkości dokonywania interupcji w błędnych wypowiedziach. Można oczekiwać, że osoby dorosłe, kiedy tylko dążyć będą do jak największej poprawności, istotnie szybciej niż badani z młodszych grup wiekowych będą przerywać błędną wypowiedź w celu wprowadzenia korekty. Zastosowanie tej samej procedury badawczej (zadania powtarzania zdań i poprawiania w nich błędów wynikających z niezgodności między treścią zdania a treścią prezentowanego obrazka, zob. 5.4.2.1.3.) w trzech grupach wiekowych (późne dzieciństwo, adolescencja, wczesna dorosłość) pozwoli ustalić zakres różnic w szybkości dokonywania interupcji między grupami.

Jednocześnie szybkość reagowania na błędy może odzwierciedlać się także w liczbie reperacji ukrytych. Zgodnie ze zmodyfikowaną zasadą przerywania mówienia (Hartsuiker i

Kolk, 2001) wykrycie błędu przez monitor uruchamia dwa równoległe procesy zatrzymywania artykulacji oraz planowania korekty. Jeśli więc monitorowanie wewnętrzne wykryłoby błąd na tyle wcześnie, żeby przerwać artykulację jeszcze przed wypowiedzeniem błędu, to równoległe mogłoby zachodzić formułowanie korekty i jej wprowadzenie do bufora artykulacyjnego. Dzięki temu, po zatrzymaniu mowy mogłoby od razu dojść do reperacji w postaci odtworzenia już zrealizowanej (poprawnej) części wypowiedzi, uzupełnionej o wprowadzoną korektę w planie przedartykulacyjnym. Inną możliwą strategią reperacji wewnętrznej byłoby opóźnienie polegające na spowolnieniu lub zatrzymaniu artykulacji i – po krótkiej przerwie – wznowieniu jej (restart bez odtwarzania) w miejscu przerwania, kontynuując wypowiedź bez zdradzenia, że mogło dojść do pomyłki (Postma i Kolk, 1993). Model podwójnej pętli percepcyjnej w opracowaniu Hartsuikera i Kolka (2001) przewiduje również, że największą szansę na reperację przed uzewnętrznieniem mają końcowe części przechowywanej w buforze artykulacyjnym wypowiedzi (licząc od początku według kolejności artykulacji). Dzieje się tak, gdyż w przypadku końcowych części buforowanej frazy monitorowanie wewnętrzne ma do dyspozycji więcej czasu na analizę, dzięki czemu może zainicjować przeformułowanie tej części planu fonetycznego, zanim zostanie on zrealizowany. Przesłanka ta jest ważna dla opracowania warunków eksperymentalnych sprzyjających reperacjom wewnętrznym.

Na podstawie literatury przedmiotu można wnioskować, że stosowanie reperacji ukrytych jest bardziej zaawansowanym sposobem autokorekt, gdyż w przeciwieństwie do reperacji jawnych opierają się one wyłącznie na wewnętrznej pętli monitorowania, która – aby działać skutecznie – wymaga zaangażowania większej puli zasobów niż pętla zewnętrzna (Hartsuiker i in., 2005). O ile reperacje jawne stwierdzono już u dzieci w 2. roku życia (Clark, 1978; Laakso, 2010), o tyle nie ustalono kiedy stają one się zdolne do dokonywania reperacji ukrytych, gdyż w badaniach spontanicznych wypowiedzi można jedynie pośrednio wnioskować o tym, czy daną niepełność mowy można zaklasyfikować jako przejaw ukrytej reperacji (por. Postma, 2000). Najczęściej wskazuje się na to, że powtórzenia<sup>138</sup> sygnalizują dokonanie korekty błędu przed jego wyartykułowaniem (Levelt 1983; 1989; Postma i Kolk, 1993). Przykładowo w badaniu Manfry i in. (2016) zauważyć można interesujący trend rozwojowy w proporcjach między reperacjami przez zastąpienie (*replacements*), czyli reperacjami jawnymi oraz reperacjami przez powtórzenie (*repetitions*), utożsamianymi z

---

138 Klasyfikowanie wszystkich powtórzeń jako reperacji ukrytych (Postma i Kolk, 1993) jest o tyle kontrowersyjne, że – jak zauważają Blackmer i Mitton (1991) – powtórzenia mogą być nie tyle efektem wykrycia i korekty błędu w planie przedartykulacyjnym, co raczej wynikiem opóźnienia w przekazaniu do bufora artykulacyjnego dalszej części planu fonetycznego, co skutkuje odtworzeniem planu, który już został zrealizowany (por. Brocklehurst, 2008).

reperacjami ukrytymi. Te pierwsze są bardziej liczne w grupie 3-latków w porównaniu z grupą 4-latków, natomiast odwrotny stosunek odnotowany jest w przypadku powtórzeń, które są mniej liczne u młodszych dzieci w porównaniu ze starszymi. Przyjmując, że powtórzenia sygnalizują dokonanie korekty błędu (lub innej nieścisłości) przed jego wyartykułowaniem można uznać, że uchwycona tu została zmiana rozwojowa, pokazująca większe zaangażowanie pętli wewnętrznej i wzrost reperacji ukrytych względem jawnych. Jednocześnie należy pamiętać, że jeszcze 5-latki nie osiągają sprawności reperacji jawnych błędów, typowej dla osób dorosłych (Jeagger, 2005). Dodatkowo, biorąc pod uwagę rozwój monitorowania wewnętrznego obserwowany w okresie szkolnym (Sasisekaran i Weber-Fox, 2012), można spodziewać się, że podobny trend rozwojowy związany ze zwiększaniem się odsetka reperacji ukrytych względem reperacji jawnych da się zaobserwować na przestrzeni późnego dzieciństwa i adolescencji aż do okresu dorosłości, analogicznie do ogólnego rozwoju kompetencji językowych (Kurcz, 2005).

Przy zaproponowanej procedurze badawczej szczególna uwaga zwrócona jest w kierunku reperacji ukrytych dokonywanych poprzez strategię odroczenia (Postma i Kolk, 1993), gdyż w najmniejszym stopniu zaburzają one płynność wypowiedzi, co każe przypuszczać, że ich stosowanie wymaga zaawansowanej kontroli i koordynacji różnych poziomów produkcji mowy (spowolnienie artykulacji na czas selekcji leksykalnej poprawnego wyrazu i jego wprowadzenia do bufora artykulacyjnego). Badanie przeprowadzone w ramach niniejszego projektu pozwoli rozstrzygnąć czy (i w jakim stopniu) dzieci w wieku szkolnym, adolescenti oraz dorośli różnią się pod względem stosowania reperacji ukrytych, szczególnie stosowania strategii odraczania.

#### **5.2.2.2 Grupa hipotez dotyczących pytań badawczych 2.1.1-2.3.2**

Hipotezy z tej grupy przedstawiają zakładane związki między kontrolą procesu mówienia a poziomem funkcji wykonawczych.

#### **Hipoteza 2.1. Poziom kontroli selekcji leksykalnej ma związek z poziomem funkcji wykonawczych.**

Hipoteza 2.1.1. Odporność selekcji leksykalnej na dystrakcję semantyczną zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w okresie od późnego dzieciństwa przez adolescencję do wcześniej dorosłości.

Hipoteza 2.1.2. Wielkość efektu interferencji semantycznej maleje wraz ze wzrostem poziomu funkcji wykonawczych u osób w okresie od późnego dzieciństwa przez adolescencję do wcześniej dorosłości.

### **Uzasadnienie:**

Modele selekcji leksykalnej przyjmują różne mechanizmy wyboru wyrazów ze słownika umysłowego. Teoria Levelta (1989; 1999b; Levelt i in., 1999) zakłada, że między jednostkami leksykalnymi zachodzi swego rodzaju konkurencja, której wielkość odpowiada za czas dostępu leksykalnego oraz może tłumaczyć powstawanie niektórych błędów leksykalnych. Manipulacja dystrakcją w procedurze nazywania obrazków może nasilać tę konkurencję, powodując opóźnienia w czasie selekcji leksykalnej (efekt interferencji semantycznej) lub przeciwnie, działać facylitująco na ten proces (priming semantyczny, kiedy dystraktor występuje odpowiednio wcześniej przed obrazkiem, torując dostęp do odpowiedniej kategorii semantycznej) (Glaser i Dünghof, 1984). Manipulacja dystrakcją może wpływać też na kodowanie morfofonologiczne powodując facylitację fonologiczną, kiedy dystraktor jest powiązany formalnie a nie semantycznie z nazwą obrazka i występuje po wyświetleniu obrazka (Schriefers i in., 1990). Aby przeciwdziałać dystrakcji ze strony kontekstu, w jakim odbywa się mówienie, proces produkcji mowy jest wyposażony w mechanizmy kontroli „górną-dół”, które weryfikują, czy aktywowany podczas selekcji leksykalnej węzeł (np. *lemma*) jest zgodny z węzłem na poziomie nadrzędnym (np. pojęcie leksykalne), który został aktywowany zgodnie z intencją nadawcy (np. rozpoznanie obiektu na obrazku, który ma być nazwany), a nie np. w wyniku dystrakcji (Levelt i in., 1999).

Według Roelofsa (2005b; 2020a) mechanizmy weryfikacji działające podczas selekcji leksykalnej, podobnie jak inne procesy kontrolne mówienia (monitorowanie też korzysta z tych mechanizmów) wykazują związki z kontrolą wykonawczą, angażując w dużej mierze te same obszary korowe i podkorowe mózgu. Badanie empiryczne współprowadzone przez Roelofsa (Shao i in., 2012) wykazało, że czas reakcji werbalnej w zadaniu nazywania przedmiotów oraz czynności na obrazkach istotnie negatywnie koreluje z miarą kontroli hamowania (zadanie sygnału stop) oraz czas reakcji werbalnej w zadaniu nazywania czynności na obrazkach istotnie negatywnie koreluje z miarą pamięci roboczej. Wynik ten pokazuje więc, że proces mówienia, w swojej najbardziej podstawowej postaci (produkowanie pojedynczych wyrazów) podlega kontroli, przynajmniej częściowo wspólnej dla innych czynności intencjonalnych. Nie rozstrzyga jednak tego, czy kontroli tej podlega selekcja leksykalna, czy któryś z innych poziomów generowania mowy, zwłaszcza, że nie manipulowano w badaniu

kontekstem nazywania i nie stosowano zewnętrznej dystrakcji językowej. Dalszych rozstrzygnięć dostarcza kolejne badanie wykonane przez ten sam zespół naukowców (Shao i in., 2013), w którym wykazali oni, że czas reakcji werbalnej w zadaniu nazywania obrazków z dystrakcją koreluje pozytywnie z czasem reakcji w zadaniu sygnału stop, mierzącym poziom hamowania nieselektywnego. Według tych autorów, z racji konkurencyjnego charakteru selekcji leksykalnej (szczególnie pod wpływem dystrakcji) wykazany związek powinien być jeszcze większy przy zastosowaniu miar hamowania selektywnego (zob. rozdział 3.). Podobne rezultaty uzyskano także w zadaniu opisywania obrazków, wykazując że uczestnicy badania o wyższych wskaźnikach kontroli hamowania (selektywnego i nieselektywnego) byli mniej podatni na dystrakcję, uzyskując niższe czasy reakcji werbalnych (Sikora i in., 2016).

Na podstawie tych doniesień można oczekiwać pozytywnego związku między poziomem hamowania reakcji, a skutecznym radzeniem sobie z dystrakcją semantyczną w zadaniu nazywania obrazków, wykorzystanym w niniejszym projekcie. W związku z tym, że zakładany jest wpływ dystrakcji semantycznej zarówno na czas selekcji leksykalnej (zob. hipoteza 1.1.5.), jak i na jej skuteczność (tj. poprawność nazywania; zob. hipoteza 1.1.2.) postawione zostały hipotezy o pozytywnym związku między poziomem funkcji wykonawczych (szczególnie selektywną kontrolą hamowania), a odpornością na dystrakcję semantyczną oraz negatywnym związku między poziomem funkcji wykonawczych (szczególnie selektywną kontrolą hamowania), a wielkością efektu interferencji semantycznej.

## **Hipoteza 2.2. Efektywność automonitorowania procesu mówienia ma związek z poziomem funkcji wykonawczych.**

Hipoteza 2.2.1. Dokładność monitorowania wewnętrznego zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w okresie od późnego dzieciństwa przez adolescencję do wcześniejszej dorosłości.

Hipoteza 2.2.2. Szybkość monitorowania wewnętrznego zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w okresie od późnego dzieciństwa przez adolescencję do wcześniejszej dorosłości.

### **Uzasadnienie:**

Bezpośrednie relacje między różnicami indywidualnymi w poziomie funkcji wykonawczych i sprawności monitorowania elementów subleksykalnych w generowanej mowie nie były dotąd

przedmiotem systematycznych badań. Podstawą do sformułowania hipotez – oprócz przesłanek teoretycznych (Roelofs, 2020a) – są także wyniki raportowane w podstawowym badaniu Wheeldon i Levelta (1995). Uzyskany przez nich wynik świadczy nie tylko o tym, że pętla wewnętrznego monitorowania operuje na poziomie kodu fonologicznego (wyraz fonologiczny), ale również, że monitorowanie wewnętrzne wymaga zasobów kontroli wykonawczej. Zgodnie z przewidywaniami autorów dodatkowe zadanie odliczania podczas wykrywania fonemów w generowanych wyrazach nie spowodowało supresji wewnętrznej pętli monitorowania, gdyż głośne odliczanie (zaangażowanie bufora artykulacyjnego) oraz monitorowanie wewnętrzne (inspekcja wyrazów fonologicznych) przebiegają na odrębnych poziomach przetwarzania języka i wykorzystują różny rodzaj reprezentacji. Efektem ubocznym zastosowanej procedury zadania podwójnego (por. Piotrowski i in., 2009) był jednak niewielki spadek poprawności rozpoznań fonemów oraz opóźnienie czasu reakcji, względem średnich wyników w grupie osób badanych monitorujących swoją mowę bez dodatkowego odliczania na głos (Levelt i Wheeldon, 1995). Wynik ten pozwala na interpretację w kategoriach zasobów pamięci roboczej, która została w jakimś stopniu obciążona zadaniem odliczania, przez co pogorszyło się także wykonanie w zadaniu monitorowania fonemów. Idąc tym tropem można przypuszczać, że osoby charakteryzujące się wyższym poziomem działania pamięci roboczej będą w stanie bardziej sprawnie monitorować swoje wypowiedzi niż osoby o niższym poziomie zasobów pamięciowych. Niewykluczone, że w grę mogą wchodzić także inne składowe funkcje wykonawczych (przypuszczalnie w większym stopniu hamowanie niż giętkość poznawcza), które są potrzebne w działaniach wymagających skupienia uwagi i zintegrowania innych procesów poznawczych (Diamond, 2013; Józefacka-Szram, 2014).

### **Hipoteza 2.3. Sposób dokonywania reperacji ma związek z poziomem funkcji wykonawczych.**

Hipoteza 2.3.1. Szybkość dokonywania interupcji błędnych wypowiedzi zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w okresie od późnego dzieciństwa przez adolescencję do wcześniej dorosłości.

Hipoteza 2.3.2. Częstość stosowania reperacji wewnętrznych przez strategię odraczania zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w okresie od późnego dzieciństwa przez adolescencję do wcześniej dorosłości.

### **Uzasadnienie hipotezy:**

Jak wskazywano wcześniej, szybkość dokonywania interupcji może zależeć od wielu czynników i nie musi być natychmiastowa po wykryciu błędu, szczególnie jeśli nadawca preferuje płynność mówienia ponad jego poprawnością (Blackmer i Mitton, 1991; Seyfeddinipur i in., 2008). Jednocześnie podniesienie priorytetu poprawności (np. poprzez manipulację eksperymentalną) powinno skłaniać osoby mówiące do jak najszybszego reagowania na zakłócenia i błędy w mowie (Postma i Kolk, 1990, za: Postma i Oomen, 2005; por. Vasić i Wijnen, 2005). Tym samym można spodziewać się, że szybkość dokonywania interupcji – przy nastawieniu na maksymalizację poprawności i korygowanie wszystkich błędów – będzie różnicować osoby nie tylko ze względu na ich wiek (zob. hipoteza 1.3.1.), ale też zdolność hamowania reakcji. Inaczej mówiąc hipoteza 2.3.1 głosi, że szybkość dokonywania interupcji może współzależać od funkcji wykonawczych, w szczególności od funkcji hamowania.

Nie dysponujemy jak dotąd zbyt wieloma danymi, które pozwoliłyby zweryfikować tę hipotezę, jednak interesujących wyników dostarcza badanie van den Wildenberga i Christoffelsa (2010), którzy odnotowali wysoką korelację czasów reakcji między odpowiedziami udzielanymi werbalnie i manualnie w wykonaniu zadania sygnału stop. Badacze ci wykazali również, że osoby charakteryzujące się „dysfunkcyjną impulsywnością” mają większe problemy z wyhamowaniem reakcji werbalnych niż osoby cechujące się bardziej adaptacyjną formą szybkiego reagowania („funkcjonalna impulsywność”). Choć nie ma to bezpośredniego przełożenia na relacje między funkcjami wykonawczymi a dokonywaniem interupcji w wyniku wykrycia błędów, jest istotną przesłanką, aby podjąć się sprawdzenia tej hipotezy, zwłaszcza że projektowane badanie obejmuje osoby w 3 grupach wiekowych, u których hamowanie reakcji oraz inne funkcje wykonawcze powinny znacznie różnić się ze względu na intensywne zmiany rozwojowe.

Zaprojektowana procedura badawcza (zob. sekcja 5.4.2.1.3.) umożliwi również osobom badanym dokonywanie w części przypadków (warunek błędnego dopełnienia) reperacji ukrytych, gdyż „błędem” (niezgodność między dopełnieniem w powtarzonym zdaniu a przedmiotem używanym przez wykonawcę czynności na obrazku) obarczony jest ostatni wyraz w zdaniu, pełniący funkcję dopełnienia. Zgodnie z teorią pętli percepcyjnej (w opracowaniu Hartsuikera i Kolka, 2001) na podleganie reperacji ukrytej mają największą „szansę” elementy znajdujące się na końcu wypowiedzi przechowywanej w buforze artykulacyjnym, gdyż im więcej czasu pętla monitorowania wewnętrznego ma na analizę buforowanego tekstu, tym większa szansa na podjęcie działań naprawczych w planie

przedartykulacyjnym. Można więc spodziewać się, że do pewnego stopnia o możliwości reperacji ukrytych decydować może pojemność bufora artykulacyjnego oraz sprawność z jaką pętla wewnętrzna monitoruje przechowywany w nim materiał. Co więcej, Hartsuiker, Kolk i Martensen (2005) wnioskuje, że wewnętrzny kanał monitorowania konsumuje więcej zasobów niż kanał zewnętrzny. Innymi słowy, uważają oni, że monitorowanie mowy wewnętrznej, służąc prewencyjnemu zapobieganiu błędów, wymaga większego wysiłku (*effort*), niż monitorowanie zewnętrzne, które może jedynie wykryć błąd już po jego wypowiedzeniu, przez co nie jest w stanie uchronić płynności wypowiedzi przed niepożądanymi zakłóceniami (Hartsuiker, 2007). Jednocześnie reperacje ukryte dokonywane przez strategię odtwarzania, chronią co prawda wypowiedź przed wystąpieniem błędu, ale – podobnie jak reperacje jawne – polegają na zatrzymaniu mowy i odtworzeniu części już zrealizowanego planu, przez co także zakłócają płynność mowy (Postma i Kolk, 1993). Mogą więc nasilić się, kiedy priorytetem jest poprawność, ale niekoniecznie, kiedy skupienie (*focus*) monitora ukierunkowane jest na inne kryteria (Vasić i Wijnen, 2005) lub kiedy liczy się przede wszystkim płynność mowy. Inaczej jest w przypadku strategii odroczenia, która polega na zwolnieniu lub zatrzymaniu artykulacji, a następnie wznowieniu jej od momentu zawieszenia (lub spowolnienia) bez jakiegokolwiek odtwarzania już zrealizowanego planu. Tego typu korygowanie błędów nie tylko chroni poprawność wypowiedzi, ale też w najmniejszym stopniu zakłóca płynność artykulacji. Wydaje się więc, że reperacje przez odroczenie pozwalają zachować równowagę między poprawnością i płynnością. Z tego względu reperacje przez odraczenie mogą być najbardziej wymagającymi korektami pod względem zaangażowania zasobów poznawczych i kontroli wykonawczej. Równoległa koordynacja dwóch ostatnich „pięter” produkcji mowy przez dopasowanie tempa artykulacji bieżącej części wypowiedzi do jednocześnie prowadzonej rewizji buforowanego planu fonetycznego wydaje się trudniejsze niż sekwencyjne działanie naprawcze w postaci interupcji, edytowania korekty i następnie zainicjowania reperacji wypowiedzi. Tym samym częstość stosowania różnych strategii może być pokłosiem dostępnych zasobów poznawczych, które można zaangażować do operacji związanych z tymi strategiami. Innymi słowy zdolność do reperacji ukrytych, a szczególnie reperacji dokonywanych poprzez odroczenie może być zależna od posiadanych zasobów definiowanych w kategoriach funkcji wykonawczych.

Interesujące w tym kontekście są wyniki badania Engelhardta i in. (2013), którzy wykazali, że większa liczba jawnych reperacji oraz powtórzeń (badacze nie klasyfikowali powtórzeń jako reperacji ukrytych) związana była z niższym poziomem hamowania reakcji

(liczba reperacji negatywnie korelowała także z miarą giętkości poznawczej). Niestety badacze w analizie statystycznej nie rozróżniali reperacji błędów leksykalnych lub fonologicznych od rewizji szyku zdania i wymiany jego składników (ten drugi rodzaj korekt stanowił ~60% wszystkich odnotowanych reperacji), przez co trudno stwierdzić, czy większa liczba reperacji przy niższym poziomie hamowania dotyczy wszystkich korekt, czy któregoś z jej typów. Ponadto brak uwzględnienia reperacji ukrytych (powtórzenia i pauzy liczone były jako osobne kategorie niepłynności) nie pozwala rozstrzygnąć, czy osoby, które miały wyższe wskaźniki hamowania poprawiały mniej swoich wypowiedzi niż osoby ze słabszym poziomem hamowania, czy raczej robiły to bardziej płynnie dokonując większej liczby reperacji ukrytych<sup>139</sup>. Tym bardziej, jeśli wykazano negatywne korelacje między częstością reperacji jawnych a poziomem hamowania oraz giętkości poznawczej warto podjąć się odpowiedzi na pytanie czy taka (lub odwrotna) relacja może zachodzić między częstością reperacji ukrytych a poziomem funkcji wykonawczych.

## 5.3 Struktura badanych zmiennych oraz operacjonalizacja

### 5.3.1 Definicje i operacjonalizacja zmiennych

Zgodnie z przyjętym modelem teoretycznym i wyróżnionymi na jego podstawie momentami powstawania, monitorowania i korygowania leksykalnych błędów mowy (zob. część 5.1.) przyjęto następujące zmienne oraz ich definicje i operacjonalizację.

**Kontrola procesu mówienia (Y)** – jest to łączne określenie na ogół hipotetycznych mechanizmów i procesów odpowiedzialnych za organizację przebiegu produkcji mowy na jej wszystkich poziomach (od konceptualizacji aż po artykulację oraz słuchowy odbiór własnej mowy). Kontrola procesu mówienia zapewnia realizację intencji nadawcy poprzez nadzór nad sprawnym planowaniem wypowiedzi oraz programowaniem kodowania językowego, a także przez monitorowanie przygotowanego tekstu w celu wykluczenia błędów i nieścisłości w komunikacie kierowanym do odbiorcy.

Na potrzeby niniejszej rozprawy wybrano trzy kluczowe aspekty zmienności kontroli procesu mówienia:

**Kontrola selekcji leksykalnej (Y.1.)** – poziom efektywności z jaką procesu produkcji mowy dokonuje wyboru odpowiedniej jednostki leksykalnej ze słownika umysłowego, blokując tym samym jednostki konkurencyjne, zarówno w warunkach typowego generowania wyrazów

---

<sup>139</sup> Inne ograniczenie tego badania wynika z faktu, że analizowano dane adolescentów i dorosłych bez kontroli wieku ani podziału na grupy wiekowe. Nie można wykluczyć, że obserwowane zależności wynikają z innych czynników rozwojowych niż dojrzewanie funkcji wykonawczych.

(nazywanie przedmiotów na obrazkach), jak i w warunkach eksperymentalnie wprowadzonej dystrakcji. Kontrola selekcji leksykalnej obejmuje zatem aspekty skuteczności selekcji leksykalnej, jej odporności na dystrakcję, jak i czasu potrzebnego na osiągnięcie dostępu leksykalnego (szybkość selekcji leksykalnej).

**Monitorowanie wewnętrzne (automonitorowanie) (Y.2.)** – proces dokonywania przez hipotetyczny monitor inspekcji wewnętrznego planu mowy i wykrywania w nim nieprawidłowości (w przypadku błędów mowy) oraz elementy struktury subleksykalnej (np. sylaby lub fonemy w przypadku zadania eksperymentalnego). Na podstawie dokładności monitorowania (stopień w jakim monitorowanie prowadzi do trafnych rozpoznań) i czasu jego trwania (szybkość dokonywanych rozpoznań) oceniana jest **efektywność monitorowania wewnętrznego**.

**Dokonywanie reperacji (Y.3.)** – sposób w jaki podejmowana jest reakcja na wykryty błąd (niezgodność treści wypowiedzi z jej referencją) poprzez zatrzymanie mówienia (interrupcja) i jawne skorygowanie reparandum lub dokonanie reperacji ukrytej.

Dla każdej z wyróżnionych zmiennych zależnych (aspektów kontroli procesu mówienia) wyróżniono kilka poziomów oraz zdefiniowano ich wskaźniki, które przedstawia tabela 3.

Zgodnie z przyjętym celem pracy wyróżniono także dwie główne zmienne niezależne podmiotowe, tj. wiek oraz poziom funkcji wykonawczych, które zdefiniowane zostały w tabeli 4. Aby zweryfikować przyjęte hipotezy zaplanowano także manipulację eksperymentalną w zadaniach psycholingwistycznych, polegającą na kontroli kontekstu semantycznego (zadanie *Nazywania obrazków*), kontroli struktury monitorowanych wyrazów (zadanie *Monitorowania sylab*) oraz określeniu miejsca korygowanego błędu w strukturze zdania (zadanie *Powtarzania zdań*). Definicja oraz operacjonalizacja tych zmiennych zamieszczona została w tabelach (tabela 4 i tabela 5).

Oprócz wyróżnionych w modelu zmiennych zależnych i niezależnych, zidentyfikowano także zmienną uboczną „płeć metrykalna”, która kontrolowana była przez dobór osób badanych do grup wiekowych z dbałością o zachowanie równowagi pod względem równej liczby osób płci męskiej i żeńskiej. Szczegóły dotyczące liczebności osób ze względu na płeć przedstawiono w sekcji 5.4.1.3.

**Tabela 3***Zmienne zależne*

Zmienna	Rodzaj zmiennej	Definicja	Wskaźnik	Pomiar	
Y1. Kontrola selekcji leksykalnej	Y1.1. Skuteczność selekcji leksykalnej	ilościowa	Przebieg selekcji leksykalnej skutkujący wyborem prawidłowej jednostki leksykalnej. Skuteczność selekcji leksykalnej objawia się wypowiedzeniem wyrazu, który odnosi się do przedmiotu wypowiedzi (np. obiektu na obrazku).	Odsetek poprawnych odpowiedzi w zadaniu „Nazywanie obrazków”	Zadanie „Nazywanie obrazków”
	Y1.2. Odporność selekcji leksykalnej na dystrakcję	ilościowa	Zachowanie wysokiego poziomu skuteczności selekcji leksykalnej przy obecności zewnętrznej dystrakcji semantycznej. Im większe obniżenie skuteczności selekcji leksykalnej pod wpływem dystrakcji, tym mniejsza jej odporność na dystrakcję.	Różnica między odsetkiem poprawnych odpowiedzi w nazywaniu obrazków w próbach bez dystrakcji (warunek neutralny) a odsetkiem odpowiedzi poprawnych w próbach z dystrakcją semantyczną (warunek dystrakcji)	Zadanie „Nazywanie obrazków”
	Y1.3. Szybkość selekcji leksykalnej	ilościowa	Czas w jakim dokonywany jest wybór właściwej jednostki leksykalnej podczas mówienia (np. nazywania przedmiotu). Im szybciej przebiega selekcja leksykalna, tym szybciej następują dalsze etapy produkcji mowy: kodowanie morfo-fonologiczne, kodowanie fonetyczne oraz artykulacja.	Średni czas reakcji werbalnej (tylko próby z poprawną odpowiedzią) w zadaniu „Nazywanie obrazków”	Zadanie „Nazywanie obrazków”
	Y1.4. Wielkość efektu interferencji semantycznej	ilościowa	Spowolnienie czasu efektywnej selekcji leksykalnej w wyniku wystąpienia dystrakcji semantycznej. Wyraża się spowolnieniem szybkości nazywania obrazków w próbach, w których wprowadzona została zewnętrzna	Różnica między średnim czasem reakcji werbalnej w warunkach dystrakcji i średnim czasem reakcji werbalnej w warunkach neutralnych w zadaniu „Nazywanie obrazków”	Czas reakcji w próbach zaliczonych jako prawidłowe w

			dystrakcja semantyczna w porównaniu do prób, które nie zawierają zewnętrznej dystrakcji semantycznej.	$(EIS = Mrt_{inter} - Mrt_{neutr})$ .	zadaniu „Nazywanie obrazków”
Y2. Automonitorowanie wewnętrzne	Y2.1. Dokładność monitorowania wewnętrznego	ilościowa	Prawidłowe rozpoznanie w bezgłośnie generowanym wyrazie obecności docelowej (wskazanej w zadaniu) sylaby lub prawidłowe rozpoznanie w tym wyrazie jej braku. Np. wykrycie w wyrazie „latawiec” sylaby /la/ oraz niewykrycie w tym wyrazie sylaby /pa/, która w nim nie występuje.	Odsetek poprawnych odpowiedzi w próbach z części 2. („Nazywanie bezgłośnie”) zadania „Monitorowanie sylab”	Zadanie „Monitorowanie sylab”
	Y2.2. Szybkość monitorowania wewnętrznego	ilościowa	Czas potrzebny na efektywne monitorowanie wewnątrz wygenerowanego wyrazu i podjęcie decyzji co do obecności lub nie w jego strukturze sylaby docelowej	Średni czas reakcji (reakcje poprawne) w próbach z części 2. („Nazywanie bezgłośnie”) zadania „Monitorowanie”	Zadanie „Monitorowanie sylab”
Y3. Dokonywanie reperacji	Y3.1. Szybkość dokonywania interupcji	ilościowa	Czas potrzebny na zatrzymanie artykulacji po wykryciu błędu liczony od rozpoczęcia artykulacji błędnego wyrazu do przzerwania artykulacji w trakcie wypowiedzania zdania lub do czasu ukończenia błędnego zdania i rozpoczęcia reperacji.	Liczba sylab zdania zawierającego błąd wypowiedzianych do momentu przzerwania artykulacji lub rozpoczęcia artykulacji reperacji w zadaniu „Powtarzanie zdań” (warunek błędnego podmiotu)	Zadanie „Powtarzanie zdań”
	Y3.2. Częstość reperacji ukrytych	ilościowa	Umiejętność dokonywania reperacji w planie przedartykulacyjnym (korekta tekstu przechowywanego w buforze artykulacyjnym) skutkująca większym odsetkiem reperacji ukrytych przez strategię odroczenia (spowolniona artykulacja przed potencjalnym błędem, brak ujawnienia błędu w wypowiedzi, brak odtworzenia już wypowiedzianego tekstu)	Odsetek zdań poprawnych (bez błędu i bez zewnętrznej reperacji poprzez odtworzenie frazy lub całego zdania) rejestrowanych w zadaniu „Powtarzanie zdań” (warunek błędnego dopełnienia)	Zadanie „Powtarzanie zdań”

**Tabela 4***Zmienne niezależne podmiotowe*

Zmienna	Rodzaj zmiennej	Definicja	Wskaźnik	Pomiar	
X.1. Wiek osób badanych	X.1.1. Wiek kalendarzowy	ilościowa	Zmienna metrykalna odnosząca się do wieku uczestników badania, wyrażona w liczbie miesięcy, które upłynęły od daty urodzenia osoby badanej do dnia w którym wzięła ona udział w badaniu.	Liczba miesięcy między datą urodzenia a datą badania.	Wywiad z osobą badaną lub jej opiekunem (w przypadku dzieci)
	X.1.2. Grupa wiekowa	porządkowa	Zmienna demograficzna odnosząca się do wieku uczestników badania i związanej z tym przynależności do jednej z 3 grup: <ul style="list-style-type: none"><li>- dzieci szkolne,</li><li>- młodzież szkolna,</li><li>- młodzi dorośli.</li></ul> Grupy zostały wyróżnione na podstawie przyjętych w literaturze przedmiotu okresów rozwojowych: późne dzieciństwo, adolescencja (okres dorastania) i wczesna dorosłość oraz odpowiadających im okresów edukacyjnych: 1. etap edukacyjny (etap edukacji wczesnoszkolnej); 3. etap edukacyjny (etap edukacji gimnazjalnej); po 4. etapie edukacyjnym (etap po ukończeniu szkoły ponadpodstawowej i ewentualnego podjęcia edukacji w szkołach wyższych i policealnych).	Przypisanie każdej osoby badanej do wyłącznie jednej grupy wiekowej na podstawie wieku metrykalnego oraz uczęszczania do (lub ukończenia) danego typu szkoły: szkoła podstawowa, gimnazjum, szkoła ponadpodstawowa, szkoła policealna lub studia wyższe.	Na podstawie wieku metrykalnego uzyskanego w wywiadzie oraz sposobu rekrutacji osób badanych.

X.2. Poziom funkcji wykonawczych	X.2.1. Kontrola hamowania	ilościowa	Poziom funkcji wykonawczej kontrola (selektywna) hamowania. Wyraża się umiejętnością powstrzymania dominującej, lecz błędnej reakcji w warunkach konkurencji między możliwymi reakcjami, wynikającej z zewnętrznych dystraktorów. Pozwala na skuteczne działanie w warunkach interferencji (inaczej: kontrola interferencji).	Koszt konfliktu - spadek poprawności wykonania zadania z flankerami wynikający z zastosowania bodźców dystrakcyjnych o kierunku przeciwnym do kierunku bodźca centralnego. Koszt konfliktu wyrażany jest liczbowo w postaci różnicy między procentem poprawnych odpowiedzi w warunku z flankerami zgodnymi, a procentem poprawnych odpowiedzi w warunku z flankerami niezgodnymi.	Zadanie z flankerami
	X.2.2. Pamięć robocza	ilościowa	Poziom funkcji wykonawczej pamięć robocza w zakresie sprawności odświeżania ( <i>updating</i> ). Wyraża się umiejętnością aktualizowania w pamięci materiału podlegającego dynamicznym zmianom.	Procent poprawnych odpowiedzi w zadaniu n-wstecz.	Zadanie n-wstecz
	X.2.3. Giętkość poznawcza	ilościowa	Poziom funkcji wykonawczej giętkość poznawcza wyrażający się umiejętnością sprawnego przełączania się ( <i>shifting</i> ) między kryteriami poprawności wykonywanego zadania.	Procent błędów perseweracyjnych w zadaniu sortowania kart.	Zadanie BCST (komputerowa wersja WCST)

**Tabela 5***Zmienne niezależne manipulacyjne*

Zmienna		Rodzaj zmiennej	Definicja	Sposób kontroli	Zakres użycia
X3.1. Kontekst semantyczny	X.3.1.1 Poziom dystrakcji semantycznej	nominalna	Poziom dystrakcji semantycznej w danej próbie nazywania obrazków definiowany przez obecność lub brak zewnętrznej dystrakcji semantycznej w postaci podawanego dousznie (drogą akustyczną) wyrazu powiązanego semantycznie z nazwą obrazka. Obecność lub brak dystrakcji semantycznej tworzy dwa warunki: 1) Warunek neutralny – próby w których nie występuje dodatkowa zewnętrzna dystrakcja semantyczna. 2) Warunek dystrakcji – próby w których występuje zewnętrzna dystrakcja semantyczna.	Dwóm spośród czterech wystąpień danego obrazka w zadaniu towarzyszy dystrakcja semantyczna, tworząc 38 prób w warunku neutralnym oraz 38 prób w warunku interferencji. Wszystkie próby w zadaniu występują w porządku losowym.	Zadanie „Nazywanie obrazków”
X.3.2. Struktura tekstu	X.3.2.1 Długość wyrazu	ilościowa (dychotomiczna)	Liczba sylab z jakich składa się wyraz poddawany monitorowaniu w celu wyszukania w jego strukturze sylaby docelowej. Przyjmuje 2 wartości: 2 sylaby lub 3 sylaby	W próbach, które nie zawierają w nazwie obrazka sylaby docelowej (prawidłowa odpowiedź: NIE) wyrazy poddawane monitorowaniu składają się z 2 lub 3 sylab.	Zadanie „Monitorowanie sylab”
	X.3.2.2. Pozycja sylaby docelowej	porządkowa	Pozycja sylaby docelowej w strukturze trój sylabowego wyrazu poddanego monitorowaniu. Przyjmuje 3 wartości w warunku zgodności: nagłos, śródgłos lub wygłos.  Warunek niezgodności: brak	Wyrazy poddawane monitorowaniu zawierają sylabę docelową na 1, 2 lub 3 pozycji (nagłos, śródgłos lub wygłos) – dotyczy prób, w których nazwa obrazka zawiera sylabę docelową (prawidłowa odpowiedź: TAK)	Zadanie „Monitorowanie sylab”

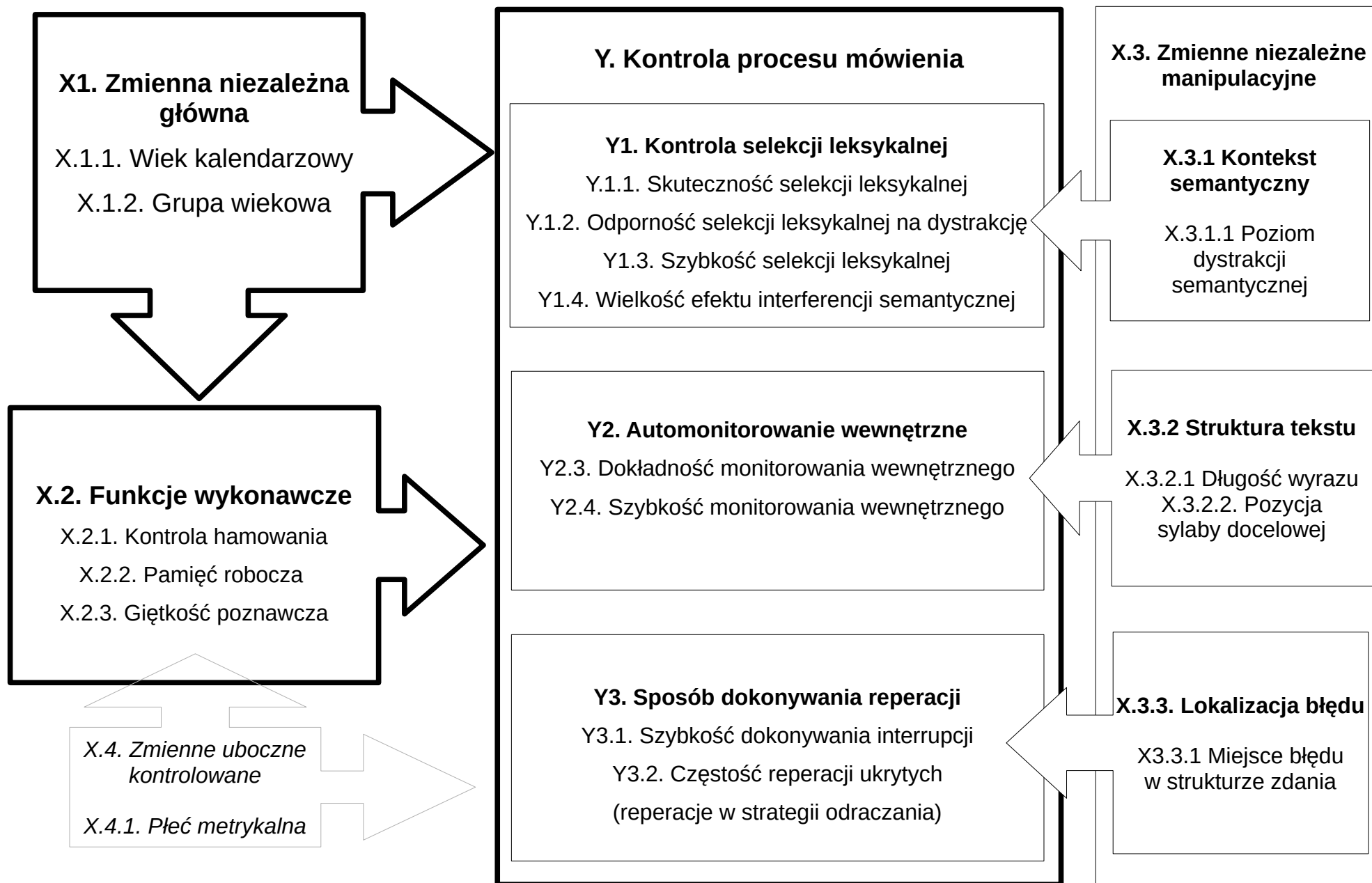
X.3.3. Lokalizacja błędu	X3.3.1 Miejsce błędu w strukturze zdania	nominalna	Przypisanie błędnej jednostki leksykalnej do jednego z dwóch rzeczowników, pełniących funkcję podmiotu lub dopełnienia w powtarzonym zdaniu o strukturze SOV. Błąd oznacza niezgodność między wypowiedzianym wyrazem (rzeczownikiem) a postacią (agens) lub przedmiotem (obiekt) na prezentowanym obrazku. Przyjmuje 3 wartości: 1) poprawny podmiot i poprawne dopełnienie (brak błędu), 2) błędny podmiot i poprawne dopełnienie lub 3) poprawny podmiot i błędne dopełnienie.	Wykorzystanie w procedurze badania prób zawierających zdanie z błędnym podmiotem (niezgodność podmiotu zdania z osobą będącą agensem czynności na obrazku) lub dopełnieniem (niezgodność dopełniania zdania z przedmiotem będącym obiektem czynności na obrazku). Warunki te były rozłączne i zawsze korekty wymagał jeden wyraz (podmiot lub dopełnienie). Oprócz warunku z błędnym podmiotem oraz warunku z błędnym dopełnieniem występował też warunek kontrolny (warunek neutralny) z brakiem błędu (treść wypowiedzianego zdania była zgodna z treścią obrazka). Wszystkie próby w zadaniu występowały w porządku losowym.	Zadanie „Powtarzanie zdań”
-----------------------------	--	-----------	--	---	----------------------------

### 5.3.2 Układ zmiennych

Wzajemne relacje opisanych powyżej zmiennych w przyjętym modelu badawczym przedstawia schemat na rysunku 9.

Rysunek 9

Układ zmiennych



## 5.4 Metoda

Badania zostały w większości sfinansowane z grantu badawczego „Popraw się!” - *Zmiany rozwojowe i poznawcze uwarunkowania funkcjonowania mechanizmów kontroli i monitorowania mowy* przyznanego przez Instytut Psychologii UAM. Na etapie przygotowania do badań (przed przystąpieniem do rekrutacji osób badanych) zasięgnięto opinii Komisji Etyki ds. Projektów Badawczych działającą przy Instytucie Psychologii UAM (obecnie Komisja Etyki ds. Projektów Badawczych Wydziału Psychologii i Kognitywistyki) nt. projektu. Przedłożony projekt otrzymał pozytywną opinię Komisji (opinia z 20.06.2016).

### 5.4.1 Uczestnicy badania

#### 5.4.1.1 Informacje wstępne

W badaniu udział wzięło 151 osób w wieku 7–27 lat ( $M=15$ ;  $SD=6$ ;  $Me=15$ ). Dobór próby miał charakter celowo-losowy, przy czym koniecznym kryterium doboru było uzyskanie świadomej zgody na udział w badaniu (dobór przez świadomą zgodę). W związku z celem badania, zakładającym porównania międzygrupowe, rekrutację do każdej z trzech grup wiekowych (dzieci, młodzież, dorośli) przeprowadzano osobno.

#### 5.4.1.2 Procedura rekrutacji

Rekrutacja osób do badania przebiegała następująco:

##### 1. Rekrutacja dzieci

W badaniu brały udział dzieci z klas 2-3 Szkoły Podstawowej Nr 1 w Poznaniu. Po uzyskaniu zgody dyrekcji szkoły na prowadzenie badań na terenie placówki osoba prowadząca badanie (autor rozprawy) wzięła udział w zebraniach z rodzicami i przedstawiła krótko projekt badawczy, jego cele oraz formę podziękowania dla dzieci, które wezmą udział w badaniu. Następnie rodzice<sup>140</sup>, którzy byli gotowi na kontakt w sprawie poznania szczegółów badania pozostawiali informację o preferowanej formie kontaktu (np. numer telefonu, adres e-mail). W kolejnym kroku osoba prowadząca badanie osobiście kontaktowała się z rodzicami, którzy pozostawili dane kontaktowe, opowiadając o sposobie przeprowadzania badania, sposobie zapewnienia bezpieczeństwa i komfortu dziecka oraz odpowiadając na pytania rodziców. W przypadku wyrażenia przez rodzica ustnej zgody na udział dziecka w badaniu

---

<sup>140</sup> Procedura rekrutacji przewidywała uzyskania zgody rodzica lub prawnego opiekuna dziecka. W przypadku wszystkich dzieci biorących udział w badaniu opiekunami prawnymi, z którymi się kontaktowano byli ich rodzice.

umawiano termin badania, sposób przekazania materiałów pisemnych nt. badania (w tym formularza pisemnej zgody na udział dziecka w badaniu) oraz organizację przybycia dziecka na badanie i tego, w jaki sposób zapewnić bezpieczeństwo dziecka po zakończonym badaniu (dzieci były odbierane bezpośrednio po badaniu przez uprawnioną osobę lub odprowadzane do szkolnej świetlicy). Po uzyskaniu pisemnej zgody rodzica na udział dziecka w badaniu, w umówionym wcześniej terminie dziecko było odbierane przez osobę prowadzącą badanie po lekcjach z klasy lub ze świetlicy szkolnej i prowadzone na miejsce na terenie szkoły, w którym odbywało się badanie. Na miejscu przedstawiano dziecku szczegóły badania oraz pytano, czy wyraża zgodę na to, aby wziąć udział w badaniu. Zapewniano również, że badanie może zostać przerwane przez uczestnika w dowolnym momencie, bez względu na powód. Informowano także o możliwym podzieleniu badania na dwie sesje oraz o dopuszczalnym robieniu przerw na zaspokojenie potrzeb fizjologicznych. Do zakończenia sesji badania nie informowano dzieci o przewidywanej formie podziękowania. Po uzyskaniu ustnej zgody dziecka na udział w badaniu przystępowano do procedury.

## **2. Rekrutacja młodzieży**

W badaniu brali udział uczniowie z klas 1-3 Gimnazjum nr 63 w Poznaniu. Po uzyskaniu zgody dyrekcji szkoły na prowadzenie badań na terenie placówki osoba prowadząca badanie (autor rozprawy) spotkała się z uczniami w klasach i krótko przedstawiła projekt badawczy, jego cele oraz formę podziękowania dla osób, które wezmą udział w badaniu. Następnie uczniowie, którzy byli gotowi na kontakt w sprawie poznania szczegółów badania pozostawiali informację o preferowanej formie kontaktu (np. numer telefonu, adres e-mail). W kolejnym kroku osoba prowadząca badanie osobiście kontaktowała się z uczniami, którzy pozostawili dane kontaktowe, opowiadając o warunkach uczestnictwa w badaniu i sposobie jego przeprowadzania, a także odpowiadając na pytania uczniów. W przypadku wyrażenia przez ucznia ustnej zgody na udział w badaniu umawiano termin badania oraz sposób przekazania materiałów pisemnych nt. badania (informacja o badaniu dla rodzica z numerem telefonu oraz adresem e-mail do kontaktu z osobą prowadzącą badanie oraz formularz pisemnej zgody rodzica na udział dziecka w badaniu). Po uzyskaniu pisemnej zgody rodzica na udział dziecka w badaniu, w umówionym wcześniej terminie uczeń samodzielnie przybywał na miejsce na terenie szkoły, w którym odbywało się badanie. Na miejscu przedstawiano uczniowi szczegóły badania oraz ponownie pytano, czy

wyraża zgodę na to, aby wziąć udział w badaniu. Zapewniano również, że badanie można przerwać w dowolnym momencie, bez względu na powód. Informowano także o możliwym podzieleniu badania na dwie sesje oraz o dopuszczalnym robieniu przerw na zaspokojenie potrzeb fizjologicznych. Po uzyskaniu ustnej i pisemnej zgody ucznia na udział w badaniu przystępowano do procedury.

### **3. Rekrutacja osób dorosłych**

W badaniu brały udział osoby dorosłe w wieku 20-27 lat. Rekrutacja odbywała się za pośrednictwem strony internetowej (<http://bresz-lab.home.amu.edu.pl/>), utworzonej specjalnie dla celów rekrutacji (obecnie nieaktywna). Na stronie znajdowała się informacja o badaniu i miejscu jego przeprowadzania (Laboratorium Badania Rozwoju i Uczenia przy Instytucie Psychologii WNS UAM<sup>141</sup>), informacje o autorze badania i dane kontaktowe (numer telefonu, adres e-mail), a także formularz za pośrednictwem, którego można było dokonać zgłoszenia do udziału w badaniu. W związku z tym, że część uczestników wzięła udział w badaniu z polecenia (lub znalazła link do badania na mediach społecznościowych znajomej osoby) dobór osób do tej grupy wiekowej w próbie miał częściowo charakter „kuli śnieżnej”.

Po otrzymaniu zgłoszenia za pośrednictwem formularza, kontaktowano się z daną osobą na podany numeru telefonu. Po wyjaśnieniu wszelkich wątpliwości i otrzymaniu w rozmowie telefonicznej ustnego potwierdzenia chęci udziału w badaniu, umawiano dogodny termin badania oraz przekazywano informację o lokalizacji miejsca badania. W dniu badania, przed rozpoczęciem procedury osoba prowadząca badanie (autor rozprawy) jeszcze raz wyjaśniała szczegóły badania oraz warunki uczestnictwa (w tym prawo do rezygnacji w każdym momencie badania bez konieczności podania przyczyny), a po wyjaśnieniu wszelkich wątpliwości i uzyskaniu pisemnej zgody, przystępowano do procedury badania.

Wszystkie materiały przekazywane osobom badanym podczas rekrutacji (informacja o badaniu, formularz zgody itd.) zostały załączone do niniejszej pracy (Załącznik A).

#### **5.4.1.3 Struktura demograficzna badanej próby**

W badaniu brało udział 51 dzieci z klasy 2-3 SP1 w Poznaniu, 50 uczniów klas 1-3 Gimnazjum nr 63 w Poznaniu oraz 50 osób dorosłych rekrutowanych za pomocą dostępnego online formularza. Z analizy danych wykluczono wyniki jednego dziecka, którego odpowiedzi werbalne nie zostały zarejestrowane w sposób pozwalający na ich analizę, stąd w

---

141 Obecnie: Laboratorium Badania Rozwoju i Uczenia przy Wydziale Psychologii i Kognitywistyki UAM

dalszym opracowaniu liczebność grupy dzieci to  $N=50$ . Wszystkie osoby dorosłe miały wykształcenie średnie lub wyższe, gdyż byli to studenci studiów 1. lub 2. stopnia oraz jednolitych studiów magisterskich na różnych kierunkach i z różnych uczelni w Poznaniu. Szczegółowe dane na temat wykształcenia osób dorosłych oraz ich aktywności zawodowej i edukacyjnej nie były zbierane.

Dobór osób badanych w poszczególnych grupach wiekowych prowadzony był z uwzględnieniem zachowania równowagi pomiędzy liczebnością osób płci męskiej i żeńskiej, aby zmienna „płeć metrykalna” nie wpływała w sposób zakłócający na porównania międzygrupowe. Charakterystykę osób badanych pod względem wieku i płci przedstawia tabela 6. W celu kontroli zmiennej ubocznej „płeć metrykalna” przeprowadzono test niezależności dwóch zmiennych (grupa wiekowa i płeć) – chi-kwadrat. Nieistotny wynik testu nie pozwala na odrzucenie hipotezy zerowej o braku związku między analizowanymi zmiennymi:  $\chi^2(1, N = 150) = 4,127; p = 0,127$ . Oznacza to, że liczebność kobiet i mężczyzn w każdej z trzech grup wiekowych można uznać za zrównoważoną, a tym samym nie ma podstaw do podejrzenia, że płeć wpływa na wyniki uzyskane przez uczestników badania w zależności od wieku. Z tego powodu zmienna „płeć” nie będzie uwzględniana w dalszych analizach.

**Tabela 6**

*Charakterystyka uczestników badania w poszczególnych grupach wiekowych*

Grupa	Wiek (w latach)	Wiek (w miesiącach)	Płeć		Razem
	$M(SD)$	$M(SD)$	Kobiety ( $n$ )	Mężczyźni ( $n$ )	$n = 150$
Dzieci	8,43 (0,35)	101,16 (4,28)	23	27	50
Młodzież	14,04 (0,88)	168,96 (10,67)	33	17	50
Dorośli	22,66 (1,61)	271,96 (19,49)	29	21	50

*Adnotacja:*  $M$  – średnia;  $SD$  – odchylenie standardowe;  $n$  – liczba obserwacji;

### **5.4.2 Procedura badania i aparatura**

Badanie prowadzone było w zacisznym pokoju, bez udziału osób postronnych w trzech różnych lokalizacjach:

1. Dzieci z klas II-III szkoły podstawowej uczestniczyły w badaniu w gabinecie psychologia i pedagoga (na terenie swojej szkoły) poza godzinami ich pracy.
2. Młodzież z klas I-III gimnazjum uczestniczyła w badaniu w gabinecie psychologia i pedagoga (na terenie swojej szkoły) poza godzinami ich pracy.
3. Dorośli uczestnicy badania brali w nim udział w Laboratorium Badania Rozwoju i Uczenia przy Instytucie Psychologii WNS UAM (obecnie LABRiU WPiK UAM).

Wybór prowadzenia badania z dziećmi i młodzieżą w ich szkołach podyktowany był tym, aby zminimalizować niedogodności udziału w badaniu (koszty dojazdu, czas dotarcia na badanie, stres związany z wykonywaniem zadań w nieznanym miejscu) dla uczniów oraz ich rodziców. Szkoła, jako miejsce znane dziecku, w którym spędza ono znaczną część dnia jest przestrzenią, w której regularnie podejmuje ono wysiłek intelektualny, dlatego udział w badaniu na terenie szkoły nie odbiegał znacznie od zadań z jakimi należy się mierzyć w tej placówce. Prowadzenie badania na terenie szkoły ograniczyło także zaburzenie rytmu dnia uczniów, gdyż część dzieci mogła wziąć udział w badaniu w czasie przebywania w świetlicy szkolnej (przed lekcjami lub tuż po nich)<sup>142</sup>. Starsi uczniowie z gimnazjum także wpisywali udział w badaniu w porządek swojego planu lekcji przed zajęciami, po nich lub między lekcjami obowiązkowymi a aktywnością pozalekcyjną oraz w dni wolne od nauki.

Przestrzeń, w której prowadzono badania we wszystkich trzech lokalizacjach zorganizowana była podobnie, aby zachować podobne standardy zbierania danych. Uczestnicy ze wszystkich grup wiekowych brali udział w badaniu siedząc przy biurku, z twarzą zwróconą w stronę drzwi wejściowych do pomieszczenia. Na biurku znajdował się komputer, na którym wyświetlane były zadania badawcze, a także klawiatura i myszka (do zbierania odpowiedzi manualnych) oraz cyfrowy rejestrator dźwięku. Podczas zbierania danych w każdym zadaniu, osoba badana siedziała sama nie wchodząc w interakcję z osobą prowadzącą badanie. Osoba prowadząca badanie po podaniu instrukcji do każdego z zadań osobno oraz upewnieniu się, że osoba badana rozumie procedurę, przemieszczała się w inne

---

<sup>142</sup> Przy rozważaniu sposobu organizacji badań starano się w maksymalnym stopniu zrealizować kluczowy postulat etyki badań z udziałem dzieci wyrażony przez Hornowską i in. (2014, s. 64) „Najważniejszym celem jest zapewnienie dziecku poczucia bezpieczeństwa i poszanowanie jego godności oraz podmiotowości bez względu na wiek. Osoby przeprowadzające badanie powinny dołożyć wszelkich starań, aby stworzyć dziecku ciepłą, przyjazną, ale rzeczową, zadaniową atmosferę w trakcie badania. Warto też zadbać o to, aby badania prowadzone były w miejscu, które dziecko zna lub wcześniej miało okazję poznać, gdzie czuje się bezpiecznie.”

miejsce w pomieszczeniu zlokalizowane z boku między stanowiskiem osoby badanej i drzwiami wejściowymi (aby pozostać w zasięgu wzroku osoby badanej (gdyby potrzebne było udzielenie jakiegokolwiek pomocy lub wyjaśnień).

Wszystkie osoby uczestniczące w badaniu postępowaly wedle tej samej procedury (wszyscy rozwiązywali te same zadania w podobnej kolejności<sup>143</sup>) i przy użyciu tego samego sprzętu elektronicznego:

1. Komputer typu laptop: Acer Aspire V3 (procesor Intel Core 2.20 GHz, 8 GB RAM, karta graficzna NVIDIA GeForce GTX 760M, ekran 17,3", system Windows 8.1.)
2. Interfejs audio: PreSonus Audiobox USB
3. Słuchawki audio (nauszne): AKG K77 Perception
4. Mikrofon dynamiczny (nagłówny): Shure WH20
5. Rejestrator dźwięku: Yamaha Pocketrak PR7
6. Klawiatura komputerowa (mechaniczna): OZON Strike Battle
7. Mysz optyczna: Hama

Zadania badawcze były prezentowane za pomocą programów DMDX (Forester i Forester, 2003) oraz PEBL (Psychology Experiment Building Language, wersja 0.14) (Mueller i Piper, 2014). Analiza nagrań dokonywana była za pomocą programu Praat (Boersma, 2001).

Przed przystąpieniem do właściwych zadań badawczych, wykonywanych przy pomocy komputera, osoby badane realizowały dwa zadania treningowe:

1. PPVT – zadanie służące badaniu uwagi. Polega na jak najszybszym reagowaniu (poprzez naciśnięcie przycisku spacji) na pojawiającą się ze zmienną częstotliwością na ekranie kropkę. Jego wykonanie pozwalało osobom badanym zaznajomić się ze sposobem wyświetlania bodźców testowych na ekranie oraz z udzielaniem odpowiedzi pod presją czasową za pomocą klawiatury.
2. Obrazki – zadanie służyło zweryfikowaniu poprawności rozpoznania przedmiotów na obrazkach wykorzystanych w zadaniach „Nazywanie obrazków” oraz „Monitorowanie sylab”. Zadanie polegało na wyświetleniu uczestnikom badania pełnego zestawu obrazków, które kolejno nazywali. W przypadku błędnej odpowiedzi (niezgodność podanej nazwy z przyjętym kluczem, por. Załącznik B/C) lub braku odpowiedzi osoba prowadząca badanie podawała oczekiwaną nazwę przedmiotu, prezentowanego na obrazku.

Przed rozpoczęciem prezentacji właściwych zadań służących zebraniu danych do analizy wykonywano:

---

143 Kolejność prezentacji zadań mogła być modyfikowana ze względu na stan psychofizyczny badanych osób (np. kaszel uniemożliwiający wykonanie zadań z odpowiedziami werbalnymi), sygnalizowaną przez nich potrzebę odpoczynku oraz organizację czasu badania z uwzględnieniem harmonogramu zajęć w szkole z podziałem na okres lekcji i przerw.

1. Dopasowanie poziomu wzbudzenia klawisza głosowego – w zależności od siły głosu (intensywności sygnału) osób badanych dokonywano obniżenia lub podwyższenia wyjściowego progu aktywizacji klawisza głosowego. Zabieg ten miał na celu optymalizację rejestracji odpowiedzi osób badanych (redukcja przypadkowego wzbudzenia klawisza głosowego przez dźwięki inne niż mowa przy jednoczesnym zapewnieniu wystarczającej czułości dla rejestracji odpowiedzi werbalnych).
2. Dopasowanie głośności dźwięku dostarczanego w słuchawkach – w zależności od zgłaszanego przez osoby badane zapotrzebowania zwiększano lub zmniejszano wyjściowy poziom głośności dźwięku prezentowanego w słuchawkach.

W kolejnej sekcji znajduje się opis zadań zastosowanych w procedurze badawczej.

### 5.4.2.1 Zadania psycholingwistyczne

#### 5.4.2.1.1 Nazywanie obrazków

Zadanie nazywania obrazków ma długą tradycję w badaniach eksperymentalnych z zakresu psychologii poznawczej i psycholingwistyki (Bock, 1996; Levelt, 1999a). W swojej podstawowej wersji polega na prezentowaniu osobom badanym obrazków i rejestrowaniu ich odpowiedzi w postaci nazw obrazków (analizie podlegają dane o poprawności oraz przede wszystkim czasie nazywania). Wartością tego paradygmatu badawczego dla badania procesu mówienia, jest wysoki poziom kontroli nad treścią wypowiedzi, przy jednoczesnym zachowaniu naturalności jej tworzenia (od rozpoznania obiektu na obrazku i aktywacji właściwego pojęcia leksykalnego, do realizacji kolejnych piętér produkcji mowy). Różne warianty tego zadania polegają na manipulacji treścią obrazków, czasem i sposobem ich prezentacji oraz kontekstem wizualnym i językowym, w jakim występują. Procedura zastosowana w niniejszym projekcie odwołuje się do paradygmatu nazywania w warunkach interferencji (dystrakcji), tzw. *picture-word interference paradigm* (Schriefers i in., 1990). W tym paradygmacie badawczym oprócz wyświetlania obrazka docelowego, prezentuje się również wyraz (dystraktor) w postaci audio (nagrana wypowiedź) lub wizualnej (zapis ortograficzny). Prezentacja dodatkowego bodźca wywołuje zazwyczaj (w zależności od manipulacji SOA<sup>144</sup>) opóźnienie w nazywaniu. W badaniu Glasera i Dünghoffa (1984) najsilniejszy efekt dystrakcji (w porównaniu do warunku neutralnego, tj. bez dystrakcji) wywoływany był przy prezentacji dodatkowego wyrazu w przedziale SOA [-300 ms;+200

---

144 SOA (ang. *stimulus-onset asynchrony*) – asynchroniczność w prezentowaniu bodźców w obrębie próby eksperymentalnej, polegająca na tym, że jeden z bodźców (S1) wyświetlany jest wcześniej od innego bodźca (S2) w tej samej próbie. SOA jest miarą używaną w badaniach eksperymentalnych określającą ilość czasu pomiędzy rozpoczęciem prezentacji bodźca S1, a początkiem prezentacji bodźca S2.

ms]. Co więcej najsilniejszy efekt opóźnienia w nazywaniu osiągnąć był kiedy dystraktor i nazwa obrazka powiązane były semantycznie, a SOA wynosiło [-100 ms;+100 ms] (por. Levelt, 1993). Podobnie w badaniu (eksperyment nr 2) Schriefersa i in.(1990) najsilniejszy efekt interferencji osiągnąć był przy SOA równym -150 ms oraz 0 ms, przy czym siła dystrakcji wyrazów powiązanych semantycznie była relatywnie wyższa od innych wyrazów (niepowiązane lub powiązane fonologicznie<sup>145</sup>) towarzyszących obrazkom do nazwania. W nowszych badaniach potwierdzono efekt interferencji dla wyrazów powiązanych semantycznie (przynależność kategorialna) lub skojarzeniowo dla SOA -150 ms (Rahman i Melinger, 2007), a także – w eksperymencie Bajo i in. (2003) – wykazano, że dystrakcja semantyczna powoduje największe opóźnienie w nazywaniu przy SOA równym -100 ms w porównaniu z mniejszą asynchronią bodźców (SOA na poziomie -75 ms oraz -50 ms).

### **Opis zadania**

Procedura została zaprogramowana przez autora rozprawy w programie DMDX (Forester i Forester, 2003), który umożliwia rejestrację czasu reakcji werbalnej za pośrednictwem klawisza głosowego (*voice key*) w postaci mikrofonu podłączonego do komputera. Dodatkowo przebieg tej części badania nagrywany był w formie audio przez rejestrator dźwięku. W zadaniu, na ekranie komputera prezentowano osobom badanym obrazki, na które należało zareagować poprzez jak najszybsze wypowiedzenie nazwy przedmiotu przedstawionego na danym obrazku. W połowie prób obrazkowi towarzyszył dystraktor w postaci wyrazu emitowanego akustycznie (nagranie głosowe odtwarzane w słuchawkach nausznych), który należało zignorować. Emisja dystraktora rozpoczynała się na 100 ms przed wyświetleniem obrazka (SOA= -100 ms). Aby nasilić efekt interferencji, do funkcji dystraktorów wybrano wyrazy powiązane semantycznie lub skojarzeniowo z nazwami poszczególnych obrazków (szczegółowy opis procedury wyboru materiału do zadań *Nazywanie obrazków* oraz *Monitorowanie sylab* znajduje się w Załączniku C). W ten sposób manipulowano w zadaniu kontekstem semantycznym, tworząc dwa warunki nazywania: (a) warunek neutralny oraz (b) warunek dystrakcji (interferencji) semantycznej. W każdym z warunków dwukrotnie prezentowano 19 obrazków, wybranych do tego zadania, co łącznie dało 76 prób składających się na zasadniczą część zadania (19 obrazków x 2 powtórzenia x 2 warunki). Wszystkie próby występowały w porządku losowym. W momencie wyświetlenia obrazka (niezależnie od warunku) rozpoczynał się pomiar czasu reakcji, trwający maksymalnie 2 sekundy. Udzielenie odpowiedzi lub upływ 2 sekund powodował przerwanie

---

<sup>145</sup> Przy SOA +150 ms wyraz powiązany fonologicznie nie wywoływał dystrakcji spowalniającej nazywanie, lecz istotnie ją przyspieszał powodując tzw. efekt *facylitacji fonologicznej* (Schriefers i in., 1990).

wyświetlania danego obrazka docelowego i przejście do nowej próby. W warunku interferencji prezentacja dystraktora rozpoczynała się na 100 ms przed obrazkiem, w trakcie 0,5-sekundowej emisji punktu fiksacji wzroku. Struktura przykładowej próby przedstawiona została na rysunku 10.

Przed przystąpieniem do procedury każda osoba badana otrzymywała szczegółową instrukcję. Zadanie główne poprzedzone było blokiem testowym (7 prób), pozwalającym zapoznać się z procedurą. Po wykonaniu prób testowych i potwierdzeniu zaznajomienia się ze sposobem rozwiązania zadania, osoby badane przystępowały do części głównej.

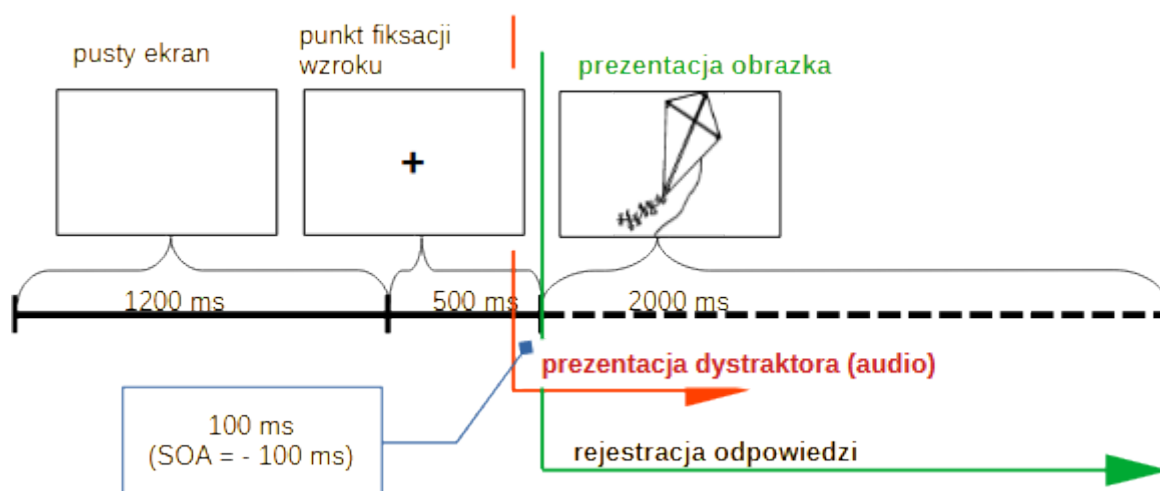
### Analiza danych

Przeprowadzona procedura dała łącznie 11 400 prób (76 prób x 150 uczestników) do analizy. W pierwszym kroku odsłuchiwano każde nagranie i oznaczano poprawność odpowiedzi na podstawie przyjętego klucza (por. Załącznik B/C). Aby wyłączyć z analizy błędy powstałe z przyczyn niezwiązanych z procedurą badania z dalszej analizy wykluczono próby:

- a) w których brak odpowiedzi lub błędna odpowiedź były wynikiem nieznamomości przedmiotu na obrazku lub jego nazwy – w tym celu zestawiano poprawność odpowiedzi w zadaniu *Nazywanie obrazków* z wcześniejszym zadaniem treningowym *Obrazki*, w którym weryfikowano znajomość przedmiotów i ich nazw, występujących na prezentowanych obrazkach;
- b) w których brak odpowiedzi lub błędna odpowiedź były wynikiem trudności powstałych we wcześniejszej (bezpośrednio poprzedzającej) próbie – w tym celu z

### Rysunek 10

Ilustracja sekwencji i czasu prezentacji bodźców w zadaniu nazywania obrazków



analizy wykluczano drugą i kolejną próbę, w których pod rząd nie padła prawidłowa odpowiedź;

- c) zakłócone czynnikami zewnętrznymi względem zadania: kaszel, kichanie, ziewanie, komentarz osoby badanej itp.

Wedle przyjętych powyżej kryteriów z analizy wykluczono 1,44% odpowiedzi. Dokładniejsze statystyki prezentuje tabela 7, w której zestawiono procent prób zakwalifikowanych do dalszej analizy.

**Tabela 7**

*Średnie i odchylenia standardowe procentu zaklasyfikowanych odpowiedzi w zadaniu nazywania obrazków*

Grupa wiekowa	Warunek		
	neutralny	dystrakcja	łącznie
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
dzieci	97,63 (4,32)	97,58 (4,25)	97,6 (4,25)
młodzież	98,37 (2,81)	98,32 (2,8)	98,34 (2,72)
dorośli	99,79 (1,04)	99,63 (1,19)	99,71 (1,07)

*Adnotacja: M – średnia; SD – odchylenie standardowe*

Sprawdzono, czy występująca między warunkami różnica (we wszystkich grupach łącznie) w procencie zaklasyfikowanych odpowiedzi jest istotna statystycznie. W tym celu przeprowadzono test Wilcozona dla par obserwacji. Uzyskano nieistotny statystycznie wynik ( $Z=-0,991$ ;  $p>0,322$ ), co oznacza, że nie ma różnicy między warunkiem neutralnym a warunkiem dystrakcji pod względem odsetka zaklasyfikowanych odpowiedzi i pozwala uznać, że proces kwalifikacji prób do analizy nie był stronniczy i nie faworyzował żadnego z warunków.

Poprawność odpowiedzi, wyznaczająca wartości zmiennej „skuteczność selekcji leksykalnej” została obliczona jedynie dla odpowiedzi zakwalifikowanych do analizy. Wskaźnik ten ustalono dla każdej osoby indywidualnie (dla każdego z warunków niezależnie) poprzez podzielenie liczby poprawnie udzielonych odpowiedzi przez liczbę odpowiedzi zaklasyfikowanych do analizy. Średni odsetek poprawnych odpowiedzi dla warunku neutralnego wyniósł  $M=97,57\%$  ( $SD=3,01$ ) a dla warunku interferencji  $M=95,36\%$  ( $SD=5,37$ ).

Porównanie poprawności między warunkami w poszczególnych grupach wiekowych jest przedmiotem szczegółowej analizy w rozdziale 6.

Program DMDX (Forester i Forester, 2003), w którym przygotowano procedurę „Nazywania obrazków” rejestrował czas reakcji w każdej próbie, a także zapisywał krótką, około 1-sekundową próbkę dźwięku (początkowy fragment wypowiedzianego wyrazu). Ze względu na to, że klawisz głosowy reagował niespecyficznym na dźwięk o określonej intensywności<sup>146</sup>, wyniki czasów reakcji automatycznie rejestrowane przez program zakłócone były przez dźwięki towarzyszące pracy aparatu artykulacyjnego, które wzbudzały klawisz głosowy przed rozpoczęciem artykulacji właściwego wyrazu. Aby uzyskać rzeczywiste czasy reakcji (szybkość rozpoczęcia nazywania obrazków) każda z prób<sup>147</sup> wymagała weryfikacji i ewentualnej ręcznej korekty zapisu czasu reakcji na podstawie próbek wypowiedzi oraz nagrania dokonanego przez rejestrator dźwięku. W Załączniku D przedstawiony został dokładny proces korygowania czasów reakcji. Ostatecznie korekty wymagało 78% zarejestrowanych prób.

Następnie, na podstawie skorygowanych czasów reakcji w odpowiedziach prawidłowych obliczono dla każdej osoby badanej średnie wartości czasu reakcji w warunku neutralnym i warunku interferencji. Uzyskano w ten sposób wartości zmiennej „czas efektywnej selekcji leksykalnej” dla obu warunków. Średnie czasy reakcji wszystkich OB w warunku neutralnym wyniosły  $M=710,45$  ms ( $SD=164,81$ ), a w warunku interferencji  $M=973,7$  ms ( $SD=282,06$ ), co dało wartość efektu interferencji na poziomie  $M=263,25$  ms ( $SD=157,84$ ). Porównanie czasów reakcji między warunkami w poszczególnych grupach wiekowych jest przedmiotem szczegółowej analizy w rozdziale 6.

#### **5.4.2.1.2 Monitorowanie sylab**

Zadanie monitorowania sylab wzorowane jest na analogicznych zadaniach opisywanych w literaturze dotyczącej badań psycholingwistycznych z udziałem osób dorosłych (Wheeldon i Levelt, 1995; Morgan i Wheeldon, 2003; Shiller, 2005). Bezpośrednią inspiracją dla zaprojektowania procedury zadania użytego w niniejszym projekcie było badanie przeprowadzone przez Sasisekaran i Weber-Fox (2012), w którym uczestniczyły dzieci w wieku 7-13 lat. Zadanie zostało więc zaprojektowane w taki sposób, aby mogły je wykonać zarówno dzieci w wieku wczesnoszkolnym, młodzież oraz osoby dorosłe, uczestniczące w badaniu.

---

146 Próg wzbudzenia klawisza głosowego dopasowywany był indywidualnie do poziomu głośności wypowiedzi osób badanych podczas instalacji sprzętu przed przystąpieniem do właściwego badania.

147 Weryfikację i korektę czasów reakcji przeprowadzono na pełnym zbiorze 11 400 prób w trakcie ich klasyfikacji i oceny poprawności odpowiedzi.

## Opis zadania

W zadaniu wyświetla się osobom badanym obrazki przedstawiające różne obiekty (zwierzęta, narzędzia, owoce i warzywa itp.). Zadaniem osób badanych jest w każdej kolejnej próbie rozpoznać obiekt na obrazku oraz odpowiedzieć, czy w jego nazwie znajduje się podana wcześniej sylaba. Jeśli w nazwie obrazka znajduje się dana sylaba należy udzielić odpowiedzi „TAK” poprzez naciśnięcie prawego przycisku SHIFT na klawiaturze komputera. W przeciwnym wypadku, tj. kiedy w nazwa obrazka nie zawiera podanej sylaby należy udzielić odpowiedzi „NIE” poprzez naciśnięcie lewego przycisku SHIFT na klawiaturze komputera. Z tego względu w zadaniu występują dwa warunki – warunek zgodności (kiedy nazwa obrazka zawiera podaną sylabę) oraz warunek niezgodności (kiedy nazwa obrazka nie zawiera podanej sylaby). Próby w obu warunkach odnoszą się do 12 sylab docelowych o strukturze CV (sylaby otwarte): /bu/, /ka/, /ko/, /la/, /le/, /mo/, /na/, /pa/, /pe/, /pu/, /ra/, /ta/.

### *Warunek zgodności*

Do warunku zgodności wybrano 14 trójsylabowych wyrazów (rzeczowników), które ilustrowane były na obrazkach pochodzących z bazy „International Picture Naming Project” (Szekely i in., 2004). Szczegółowy opis przygotowania materiału językowo-wizualnego do zadań *Monitorowanie sylab* oraz *Nazywanie obrazków* znajduje się w Załączniku C. Wyrazy (nazwy obrazków) do warunku zgodnego dobrano w taki sposób, aby każdy z wyrazów mógł wystąpić dwukrotnie (w dwóch różnych blokach) w jednej serii (całe zadanie składało się z dwóch serii po 12 bloków każda), za każdym razem w kontekście innej sylaby docelowej (np. wyraz „papuga” pierwszy raz występował w kontekście wykrywania sylaby /pa/, a po raz drugi w kontekście sylaby /pu/). Sylaba docelowa mogła występować w nagłosie (1. sylaba), śródgłosie (2. sylaba) lub w wygłosie (3. sylaba). W związku z tym, że każdy z obrazków występował dwukrotnie w jednej serii, przy każdym wystąpieniu danej nazwy obrazka, sylaba docelowa, którą należało wykryć znajdowała się na innej pozycji (zob. rysunek 11). Występowanie sylab w różnych pozycjach wyrazów nazywających obrazki zorganizowano w ten sposób, że w 10 próbach sylaba znajdowała się w nagłosie, w 9 próbach sylaba docelowa znajdowała się w śródgłosie oraz w 9 próbach sylaba docelowa znajdowała się w wygłosie. Ostatecznie więc na warunek zgodności składało się w jednej serii 28 prób (14 wyrazów zawierających sylaby docelowe w nagłosie, śródgłosie lub wygłosie). Wszystkie wyrazy przedstawia tabela 8.

**Tabela 8**

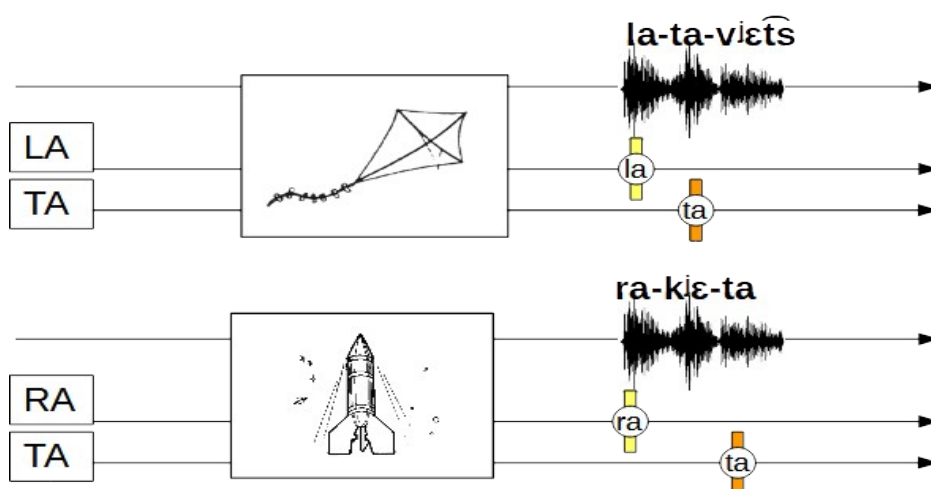
*Wyrazy (nazwy obrazków) wykorzystane w warunku zgodności ze wskazaniem miejsca artykulacji sylaby docelowej*

Nagłos		Śródgłos		Wygłos	
Zapis ortograficzny	Zapis IPA	Zapis ortograficzny	Zapis IPA	Zapis ortograficzny	Zapis IPA
parasol	/parasɔl/	parasol	/parasɔl/	kanapa	/kanapa/
papuga	/papuga/	korale	/kɔrale/	rakieta	/rakʲɛta/
rakieta	/rakʲɛta/	kapelusz	/kapɛлуʂ/	skarpeta	/skarpɛta/
kapelusz	/kapɛлуʂ/	skarpeta	/skarpɛta/	butelka	/butɛlka/
butelka	/butɛlka/	latawiec	/latavʲɛts/	motyle	/mɔtʲilɛ/
latawiec	/latavʲɛts/	cebula	/tʂɛbula/	korale	/kɔrale/
pudełko	/puɔɛwkɔ/	papuga	/papuga/	cebula	/tʂɛbula/
komoda	/kɔmɔda/	komoda	/kɔmɔda/	pudełko	/puɔɛwkɔ/
korona	/kɔrɔna/	kanapa	/kanapa/	korona	/kɔrɔna/
motyle	/mɔtʲilɛ/				

*Adnotacja: Zapis IPA wygenerowany został za pomocą konwertera <https://baltoslav.eu/>*

**Rysunek 11**

*Ilustracja miejsca artykulacji sylab docelowych w nazwach obrazków*



**Tabela 9***Wyrazy (nazwy obrazków) wykorzystane w warunku niezgodności*

Wyrazy 2-sylabowe		Wyrazy 3-sylabowe	
Zapis ortograficzny	Zapis IPA	Zapis ortograficzny	Zapis IPA
oko	ɔkɔ	koszula	kɔʂula
noga	nɔga	żarówka	ʒarufka
jabłko	japko	widelec	vʲidɛlɛʦ̥
kura	kura	drabina	drabʲina
młotek	mwoʦɛk	patelnia	patɛlʲna
zebra	zɛbra	pomidor	pɔmʲidɔr
koza	kɔza		
szklanka	ʂklanka		

*Adnotacja:* Zapis IPA wygenerowany został za pomocą konwertera <https://baltoslav.eu/>

#### *Warunek niezgodności*

Do warunku niezgodności wybrano 14 wyrazów (rzeczowników): 8 wyrazów dwusylabowych oraz 6 wyrazów trójsylabowych, które ilustrowane były na obrazkach pochodzących z tej samej bazy danych (zob. Załącznik C). Wyrazy te oraz ich dystrybucja w blokach zostały tak dobrane, aby nie zawierały sylaby docelowej, która miała być wyszukiwana w danym bloku ani difonu wchodzącego w skład tej sylaby (np. wyraz „widelec” wystąpił w bloku, w którym monitorowanie miało wykryć obecność w nazwie obrazka sylaby /ta/, ale nie sylaby /le/). Każdy z wyrazów wybranych do warunku niezgodności wystąpił dwukrotnie (28 prób). Zestawienie wyrazów wybranych do warunku niezgodności przedstawia tabela 9.

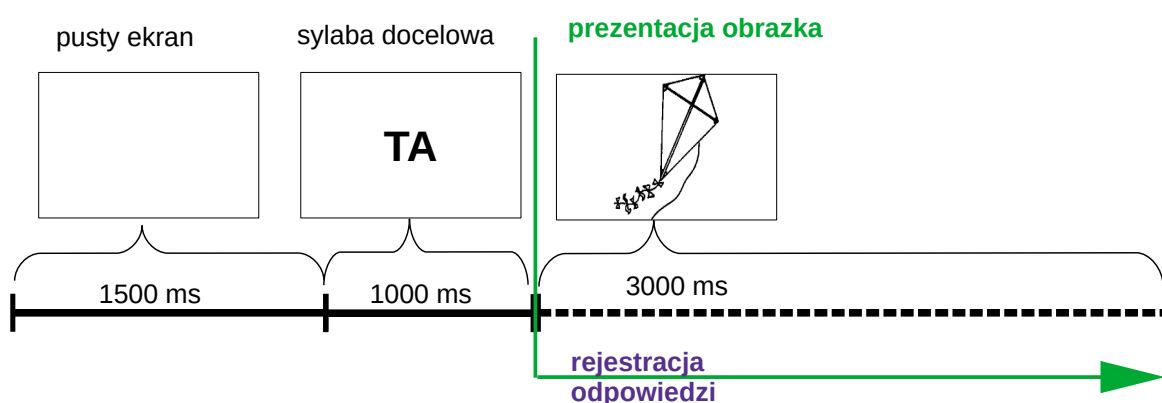
#### **Organizacja materiału**

Procedura została zaprogramowana przez autora rozprawy w programie DMDX (Forester i Forester, 2003). Próby zostały zorganizowane w dwie serie, składające się z 12 bloków każda. W pierwszej serii osoby badane nazywały prezentowane im obrazki na głos, wskazując, czy w wypowiedzianej nazwie znajduje się podana na początku danego bloku sylaba (monitorowanie zewnętrzne). Pierwsza seria została poprzedzona blokiem testowym składającym się z dodatkowych 7 prób, przedstawiających obrazki, w których nazwach

należało wyszukać sylabę /ba/. Druga seria składała się dokładnie z takich samych prób, jak seria pierwsza, jednak nazywanie obrazków odbywało się bezgłośnie (monitorowanie wewnętrzne). W obu seriach kolejność bloków z próbami była taka sama i stała dla wszystkich osób badanych, natomiast kolejność prób w każdym bloku była losowa.

## Rysunek 12

*Ilustracja sekwencji zdarzeń i czasu prezentacji bodźców w przykładowej próbie zadania Monitorowanie sylab (warunek zgodności)*



Przed każdym blokiem prezentowana była głosowo i wizualnie (zapis ortograficzny) sylaba, której należało wyszukiwać w nazwach obrazków. Następnie wyświetlane były kolejne próby (porządek losowy) w ramach danego bloku. Każda próba poprzedzona była przerwą trwającą 1,5 sekundy podczas której ekran komputera pozostawał pusty, po czym przypomniana była wizualnie (zapis ortograficzny) sylaba docelowa, której ekspozycja trwała 1 sekundę. Następnie wyświetlany był obrazek, którego nazwa miała być monitorowana pod kątem zawierania podanej sylaby. Obrazek pozostawał na ekranie do momentu udzielenia odpowiedzi przez wciśnięcie przycisku na klawiaturze komputera (lewy lub prawy SHIFT). W momencie wyświetlenia obrazka rozpoczynała się rejestracja odpowiedzi (czas reakcji), która trwała maksymalnie 3 sekundy. Po tym czasie w razie nieudzielenia odpowiedzi program automatycznie przechodził do kolejnej próby. Udzielenie odpowiedzi przed upływem 3 sekund od momentu pojawiania się na ekranie obrazka kończyło daną próbę i prowadziło do rozpoczęcia kolejnej próby. Struktura przykładowej próby przedstawiona została na rysunku 12.

Próby z bloku testowego oraz z pierwszej serii służyły opanowaniu przez osoby badane zadania oraz ustaleniu, które wyrazy należy wykluczyć z analizy ze względu na

trudności z rozpoznaniem obiektu na obrazku lub wydobyciem właściwej nazwy. Odpowiedzi z prób bloku testowego oraz pierwszej serii nie zostały uwzględnione w analizie wyników.

W drugiej serii wybór sylaby odbywał się przy bezgłośnym nazywaniu obrazków (monitorowanie wewnętrzne) i analogicznie, jak w pierwszej serii należało wskazać, czy nazwa obrazka zawiera podaną na początku bloku sylabę. Wyniki drugiej serii posłużyły do zebrania danych, które następnie opracowano, aby można było je wykorzystać w statystycznym testowaniu hipotez.

### **Opracowanie danych**

Przeprowadzona procedura dała łącznie 8 400 prób (56 prób x 150 uczestników) do analizy w każdej z obu serii (łącznie 16 800 prób). W pierwszym kroku odsłuchiowano każde nagranie pierwszej serii prób, w której osoby badane nazywały obrazki na głos. Każdą odpowiedź klasyfikowano jako poprawną lub błędną (jako błąd traktowano także brak podania nazwy) i porównywano z poprawnością nazw podawanych w zadaniu *Obrazki* (zadanie weryfikujące znajomość przedmiotów na obrazkach i ich nazw). Do dalszej analizy włączano jedynie próby drugiej serii, co do których nie było wątpliwości, że osoba badana zna i używa poprawną (zgodną z kluczem, zob. Załącznik B/C) nazwę obrazka.

W kolejnym kroku, na podstawie raportu wygenerowanego przez program DMDX, każdą zaklasyfikowaną odpowiedź kategoryzowano pod względem poprawności (poprawna vs. błędna) oraz przypisywano jej wartość czasu reakcji. Następnie dla każdej osoby badanej indywidualnie obliczano procent poprawnych odpowiedzi w poszczególnych warunkach oraz średnie czasy reakcji (wyłącznie odpowiedzi poprawne) w każdym z warunków.

Z porównań różnic w czasach reakcji, wynikających z manipulacji strukturą monitorowanego tekstu, wyłączono odpowiedzi osób, które uzyskały poprawność odpowiedzi na poziomie niższym niż 30% w którymkolwiek warunku. Z tego względu do porównań średnich czasu reakcji ze względu na pozycję sylaby docelowej w wyrazie zakwalifikowano odpowiedzi: 38 osób z grupy dzieci, 48 osób grupy młodzieży oraz 50 osób z grupy dorosłych. Analogicznie, do porównań średnich czasu reakcji ze względu na długość wyrazu zakwalifikowano odpowiedzi: 47 osób grupy dzieci, 48 osób z grupy młodzieży oraz 50 osób z grupy dorosłych (por. tabela 10).

Szczegółowe porównanie wyników w poszczególnych grupach wiekowych jest przedmiotem analizy w rozdziale 6.

**Tabela 10**

*Średnie i odchylenia standardowe procentu zaklasyfikowanych odpowiedzi w zadaniu  
Monitorowanie sylab*

Grupa wiekowa	Warunek		
	zgodności	niezgodności	łącznie
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
dzieci	94,86% (1,19)	95,5% (0,91)	95,18 % (0,98)
młodzież	97,89% (0,6)	98,83% (0,38)	98,36% (0,35)
dorośli	99,29% (0,29)	99,64% (0,26)	99,46% (0,22)

*Adnotacja: M – średnia; SD – odchylenie standardowe*

### 5.4.2.1.3 Powtarzanie zdań

Reperacje błędów mowy bada się głównie analizując spontaniczne lub eksperymentalnie indukowane przejęzyczenia. Z racji tego, że procedury eksperymentalne takie jak SLIP (Motley i in., 1982), czy powtarzanie łamańców językowych indukują błędy na poziomie subleksykalnym (lub artykulacyjnym), to eksperymentalne badanie reperacji na błędach leksykalnych opiera się w większości na wariantach zadań opisywania obrazków (np. Boland i in., 2005; Hanley i in., 2016). Zadanie „Powtarzanie zdań” przygotowano głównie w oparciu o procedurę indukowania reperacji zastosowaną przez van Wijka i Kempna (1987), polegającą na opisywaniu obrazków, których treść ulegała przekształceniom podczas mówienia o nich, prowokując konieczność skorygowania wcześniej przygotowanego planu wypowiedzi. Podstawową modyfikacją w procedurze (względem oryginału) było prezentowanie osobom badanym gotowych opisów obrazków (przed wyświetleniem obrazka) w postaci zdań do powtórzenia. Dzięki temu pominięto żmudny etap trenowania osób badanych w sposobie opisywania obrazków, który był konieczny w oryginalnej procedurze, aby zapewnić dostateczną kontrolę nad strukturą generowanej wypowiedzi. Uproszczona także została struktura zdań, ograniczając ją do szyku SVO.

### Opis zadania

W zaprojektowanym zadaniu osoby badane miały powtarzać słyszane w słuchawkach zdania, będące opisem wyświetlanych na ekranie komputera obrazków. Każda próba składała się z sekwencji:

1. Prezentacja nagrania audio zawierającego opis obrazka docelowego. Aby zachować neutralny charakter intonacji oraz równe tempo artykulacji wszystkich zdań, nagrania

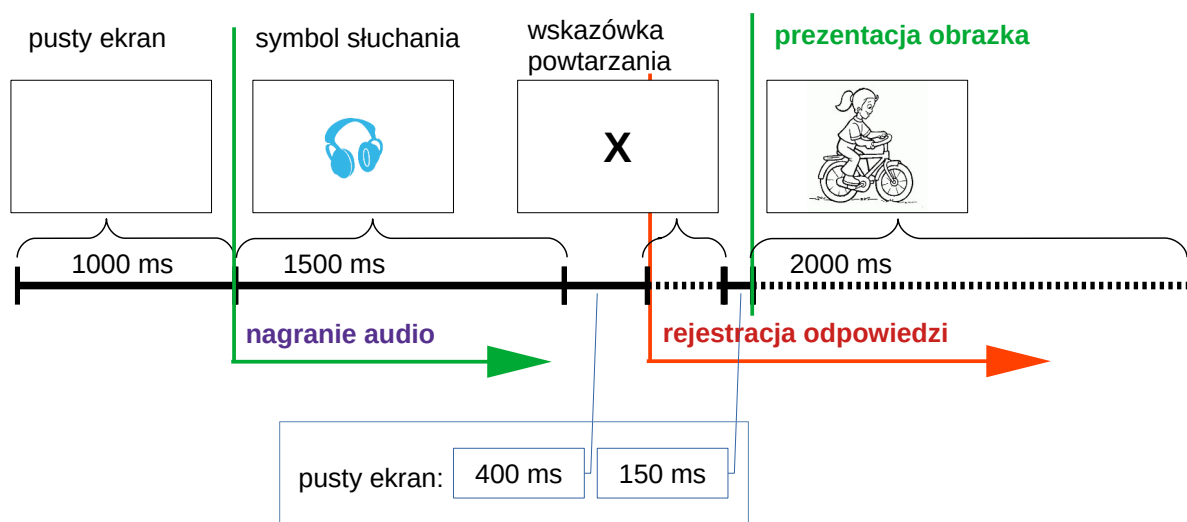
audio zostały przygotowane z pomocą syntezy mowy IVONA 1.5.1 zainstalowanego na czytniku ebooków Onyx Boox X61S i zarejestrowane w programie Audacity 2.1.2. Podczas odtwarzania nagrania na ekranie wyświetlany był symbol słuchania (ikona słuchawek), który był widoczny przez 1,5 sekundy, niezależnie od długości nagrania.

2. Sygnał do rozpoczęcia głośnego powtarzania (symbol „X” na ekranie). Sygnał pojawiał się na ekranie po 400 ms od zgaśnięcia symbolu słuchania i pozostawał wyświetlony do czasu rozpoczęcia artykulacji przez osobą badaną (odpowiedź rejestrowana za pomocą mikrofonu służącego jako klawisz głosowy).
3. Wyświetlenie obrazka docelowego. Po upływie 150 ms od zarejestrowania głosu osoby badanej, na ekranie prezentowany był obrazek docelowy, do którego odnosiło się powtarzane zdanie.

Przebieg przykładowej próby przedstawia rysunek 13. Całość procedury została zaprogramowana przez autora rozprawy w programie DMDX (Forester i Forester, 2003),

### Rysunek 13

*Ilustracja sekwencji zdarzeń i czasu prezentacji bodźców w przykładowej próbie zadania Powtarzanie zdań*



Aby sprowokować reperacje wypowiedzi osób badanych manipulowano zgodnością obrazka z treścią powtarzanego zdania. W warunkach neutralnych treść obrazka zgadzała się z treścią zdania, co znaczy że powtarzane zdanie właściwie opisywało prezentowany obrazek i osoba badana nie musiała wprowadzać żadnej korekty do swojej wypowiedzi (o ile

powtórzyła zdanie bodźcowe w takiej formie, w jakiej je usłyszała). Symulacja pomyłek słownych oraz indukowanie reperacji odbywało się poprzez wyświetlanie obrazków, które w jakimś aspekcie (wykonawca czynności lub jej przedmiot) nie zgadzały się z treścią powtarzanego zdania.

W ten sposób wprowadzone zostały dwa warunki niezgodności:

1. Warunek błędnego podmiotu. Obrazek przedstawiał innego wykonawcę czynności niż agens powtarzanego zdania, dlatego aby dopasować treść zdania do obrazka należało dokonać korekty podmiotu, znajdującego się na pierwszym miejscu w zdaniu (szyk SVO). W tym warunku konieczność dokonania naprawy ujawniana była już na samym początku wypowiedzi, co miało prowokować osoby badane, do jak najszybszej interupcji i reperacji. Błąd nie mógł zostać wykryty przed rozpoczęciem jego artykulacji, więc reperacja mogła mieć charakter jedynie jawny w postaci restartu całego zdania.
2. Warunek błędnego dopełnienia. Obrazek przedstawiał inny przedmiot czynności niż referencja dopełnienia powtarzanego zdania. Korekty wymagał zatem ostatni wyraz w zdaniu. W tym warunku błąd ujawniał się dopiero wraz z wypowiedzeniem całego zdania. W takiej sytuacji reperacja jawna mogła polegać na restarcie wypowiedzi od dowolnego momentu i wprowadzeniu na końcu właściwego wyrazu. Możliwe było też wykrycie błędu jeszcze na poziomie planu przedartykulacyjnego i korekta przed jego realizacją, co wymagało interupcji lub spowolnienia artykulacji i dokonania reperacji ukrytej.

Przykład zdania do powtórzenia oraz obrazków w trzech warunkach, a także możliwe reakcje osób badanych, przedstawione zostały na rysunku 14.

Materiał obrazkowy został przygotowany na podstawie ogólnodostępnych zasobów internetowych. Lista 20 zdań do powtórzenia wykorzystanych w zadaniu znajduje się w Załączniku E. Każde zdanie wykorzystane było dwukrotnie w każdym z warunków, co dało łącznie 60 prób. Przed przystąpieniem do badania, uczestnikom przedstawiana była szczegółowa instrukcja. Zadanie główne poprzedzone było blokiem 7 prób testowych, pozwalających osobom badanym zapoznać się z procedurą. Po wykonaniu prób testowych i potwierdzeniu zrozumienia zasady udzielania odpowiedzi, przechodzą do zadania zasadniczego.

## Rysunek 14

Przykład spodziewanych reakcji osób badanych w trzech warunkach zadania  
Powtarzanie zdań

Zdanie do powtórzenia: *Dziewczynka rysuje dom*



### Warunek neutralny

Prawidłowe zdanie (bez reperacji):

*Dziewczynka rysuje dom*



### Warunek błędnego podmiotu

1) Reperacja bez interupcji

*Dziewczynka rysuje dom... Chłopiec rysuje dom*

2) Reperacja z interupcją

*Dziewczynka rysuje d... Chłopiec rysuje dom*

*Dziewczynka rysuje... Chłopiec rysuje dom*

*Dziewczynka rys... Chłopiec rysuje dom*

*Dziewczynka... Chłopiec rysuje dom*



### Warunek błędnego dopełnienia

1) Reperacja bez interupcji

*Dziewczynka rysuje dom... Dziewczynka rysuje kwiat*

2) Reperacja z interupcją i odtworzeniem

*Dziewczynka rysuje d... Dziewczynka rysuje kwiat*

*Dziewczynka rysuje... rysuje kwiat*

*Dziewczynka rys... rysuje kwiat*

*Dziewczynka... Dziewczynka rysuje kwiat*

3) Reperacja bez odtworzenia

*Dziewczynka rysuje... kwiat*

*Dziewczynka... rysuje kwiat*

*Dziewczynka ryssssuje kwiat*

Łącznie, zastosowana procedura dała 9 000 zdań (60 prób x 150 uczestników) do analizy. Opracowanie danych rozpoczęto od przesłuchania nagrań wszystkich odpowiedzi oraz zaklasyfikowania ich i wyłączenia z dalszej analizy tych, które:

- a) były odpowiedziami błędnymi lub niepełnymi – ostateczna treść wypowiedzianego zdania (z uwzględnieniem reperacji) nie zgadzała się z treścią obrazka lub nie miała struktury SVO;
- b) zawierały zmieniony czasownik (orzeczenie), mimo iż żaden z warunków nie przewidywał korekty orzeczenia zdania wyjściowego (zdanie do powtórzenia oraz zdanie po korekcie powinny zawierać ten sam czasownik);

- c) nie zostały udzielone w przewidzianym czasie, tj. jeśli ich artykulacja nie rozpoczęła się po wyświetleniu wskazówki powtarzania, ale przed wyświetleniem obrazka opisywanego przez zdania:
- wykluczone zostały próby udzielone w czasie krótszym niż 10 ms od wyświetlenia wskazówki powtarzania – wyłączenie tych odpowiedzi z analizy miało na celu wykluczenie prób, w których odpowiedź była rozpoczynana przed pojawieniem się wskazówki (w przypadku prób, w których powtarzanie zdania rozpoczynane było przed pojawieniem się wskazówki powtarzania moment rejestracji czasu reakcji przypadał na już trwającą artykulację, przez co klawisz głosowy natychmiastowo rejestrował czas reakcji rzędu jednej lub kilku milisekund);
  - w warunku błędnego podmiotu wykluczone zostały odpowiedzi, które zawierały od razu prawidłowy opis obrazka, co było efektem zobaczenia obrazka przed powtórzeniem usłyszanego zdania, np. wskutek uruchomienia klawisza głosowego przypadkowym dźwiękiem przed rozpoczęciem właściwej artykulacji (dla klasyfikowania zdań w warunku neutralnym takie przypadki nie miały znaczenia, natomiast w warunku błędnego dopełnienia nie było możliwe wykluczenie takich przypadków ze względu na możliwość dokonywania reperacji ukrytych).

Tabela 11 przedstawia uśredniony procent odpowiedzi w poszczególnych warunkach, które poddano dalszej analizie statystycznej.

**Tabela 11**

*Średnie i odchylenia standardowe procenta odpowiedzi z zadania „Powtarzanie zdań” zaklasyfikowanych do dalszej analizy.*

Grupa wiekowa	Warunek			Łącznie
	neutralny	błędny podmiot	błędne dopełnienie	
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
dzieci	74,25% (13,61)	71,05% (14,37)	65,9% (13,35)	71,33% (11,24)
młodzież	86,57% (11,24)	86,9% (13,09)	80,8% (14,75)	84,57% (11,65)
dorośli	91,6% (11,35)	92,2% (11,35)	85,7% (14,14)	89,83% (11,35)

*Adnotacja: M – średnia; SD – odchylenie standardowe*

W dalszej kolejności zmierzono czas trwania każdej odpowiedzi (od rozpoczęcia artykulacji do zakończenia artykulacji reperaturacji) oraz policzono wszystkie sylaby, tworzące wypowiedź objętą zmierzonym czasem. Na bazie tych dwóch wartości („liczba sylab” oraz „czas trwania”) obliczono tempo mówienia<sup>148</sup> w każdej odpowiedzi zgodnie ze wzorem (Frydrychowicz, 1999, s. 104):

$$\text{Tempo mówienia} = \frac{\sum \text{sylab w etapie}}{\text{czas etapu}}$$

Następnie zostało obliczone średnie tempo mówienia dla każdej osoby z podziałem na warunki (por. Załącznik F). Tabela 12 przedstawia średnie tempo mówienia w warunkach neutralnych dla 3 grup wiekowych

W kolejnym kroku, każda odpowiedź z warunków błędnego podmiotu oraz błędnego dopełnienia została skategoryzowana pod względem sposobu dokonania reperaturacji, jaki został zastosowany:

- a) reperaturacja bez interupcji – zdanie z błędem zostało wypowiedziane w całości, a reperaturacja była dodawana na końcu (po wypowiedzeniu całego zdania zawierającego reperaturandum)
- b) reperaturacja z interupcją (i odtworzeniem) – zdanie było przerywane w trakcie artykulacji reperaturandum, po czym wprowadzana była reperaturacja obejmująca swym zakresem całe zdanie (oba warunki) lub czasownik z dopełnieniem albo samo dopełnienie (warunek błędnego dopełnienia)
- c) reperaturacja bez odtworzenia – wyłącznie w warunkach błędnego dopełnienia możliwe było rozpoczęcie powtarzania zdania i jego korekta w trakcie artykulacji bez jawnej reperaturacji z odtworzeniem (strategia odraczania). Do tej kategorii zaliczono wszystkie zdania z warunkach błędnego dopełnienia wypowiedziane bez błędów.

Dla każdej próby w obu warunkach błędów (niezgodności) wyznaczono także moment interupcji, na podstawie liczby sylab jakie zostały wypowiedziane w powtarzonym zdaniu przed wprowadzeniem reperaturacji (liczba sylab składających się na oryginalną wypowiedź z reperaturandum, por. Levelt, 1983). W przypadku reperaturacji klasyfikowanych jako „bez interupcji” (zdanie z błędem wypowiedziane do końca), za moment interupcji uznawano ostatnią sylabę błędnego zdania. Proporcja różnych sposobów dokonywania reperaturacji oraz szybkość dokonywania reperaturacji (wyznaczana średnim momentem interupcji) jest przedmiotem dokładnej analizy statystycznej w rozdziale 6.

---

<sup>148</sup> W literaturze przedmiotu odróżnia się „tempo mówienia” od „tempa wypowiedzi” (Michalik i in., 2018). Zastosowana procedura uwzględnia cały czas trwania wypowiedzi, łącznie z pauzami, dlatego bardziej odpowiada pojęciu „tempa wypowiedzi” (Cholewiak, 2016). Dla utrzymania spójności terminologicznej, zachowane jednak zostanie określenie „tempo mówienia” (Frydrychowicz, 1999).

**Tabela 12**

Średnie tempo mówienia w każdej z badanych grup wiekowych w warunku neutralnym zadania Powtarzania zdań.

Grupa wiekowa	Warunek neutralny
	<i>M (SD)</i>
dzieci	3,14 (0,54)
młodzież	3,99 (0,61)
dorośli	5,34 (0,55)

Adnotacja: *M* – średnia; *SD* – odchylenie standardowe

### 5.4.2.2 Zadania mierzące kontrolę wykonawczą

#### 5.4.2.2.1 BCST (WCST)

Test Sortowania Kart z Wisconsin (*Wisconsin Card Sorting Task, WCST*) jest jednym z najstarszych i najszerzej wykorzystywanych zadań badawczych i diagnostycznych funkcji kory przedczołowej (Diamond, 2013). Stosowany jest jako miara ogólna funkcji wykonawczych, a także jako narzędzie do badania pamięci roboczej (Piotrowski, Stettner i in., 2009). W badaniu Miyake i in. (2000) zastosowanie analizy zmiennych latentnych wykazało, że wyniki Testu Sortowania Kart z Wisconsin w głównej mierze „ładowane” są przez czynnik przełączania (*shifting*), czyli umiejętność najsilniej powiązaną z giętkością poznawczą (por. Fisk i Sharp, 2004). Podobnie w badaniu Huizinga i in. (2006) stwierdzono, że wyniki WCST były najsilniej przewidywane przez poziom funkcji przełączania, a pozostałe funkcje (pamięć robocza oraz hamowanie) nie miały istotnego wpływu na wynik testu. Z tego względu komputerową wersję WCST wybrano do pomiaru poziomu funkcji giętkości poznawczej.

#### Opis zadania

W projekcie skorzystano z procedury zadania Berg's Card Sorting Task-64 (BCST-64), która jest komputerową, skróconą wersją WCST przygotowaną w oprogramowaniu Psychology Experiment Building Language (PEBL), wersja 0.14 (Mueller i Piper, 2014). Wersja skrócona zawiera 64 karty i jest alternatywną wersją pełnego testu WCST z powodzeniem testowaną na uczestnikach w wieku 6-74 lata (Fox i in., 2013).

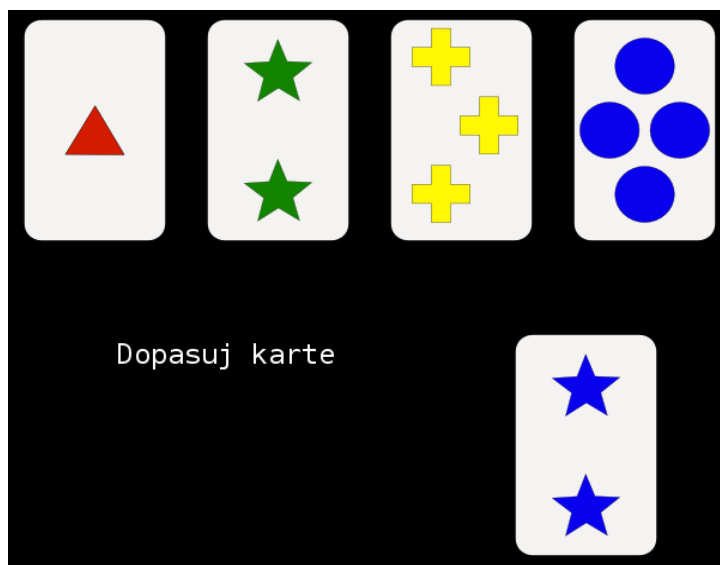
Zadaniem osób badanych było dopasowywanie kart bodźcowych (dobieranych z talii losowo) do jednej z 4 kart wzorcowych. Karty różniły się zamieszczonymi na nich kształtami

figur (gwiazdki, koła, trójkąty i krzyżyki), liczbą figur (od 1 do 4) i ich kolorem (niebieski, żółty, czerwony, zielony). Reguła sortowania nie była podana jawnie, a każdy uczestnik musiał ją odkryć samodzielnie, na podstawie otrzymywanej informacji zwrotnej o poprawności dokonanego wyboru. Informacja w postaci zielonego napisu „DOBRZE” lub czerwonego napisu „ZLE<sup>149</sup>” była wyświetlana na ekranie, po każdej reakcji wyboru. Wskazanie reguły sortowania kart docelowych dokonywane było poprzez kliknięcie lewym przyciskiem myszy komputerowej w jedną z czterech kart wzorcowych. Czas na dokonanie wyboru nie był ograniczony. Bez wiedzy osoby badanej, po 5 prawidłowo wybranych kartach następowała zmiana reguły. O zmianie reguły można było wywnioskować na podstawie informacji zwrotnej o dokonaniu błędnego wyboru.

Ekran zawierający cztery karty wzorcowe oraz przykładową kartę docelową przedstawia rysunek 15.

### Rysunek 15

*Fragment ekranu z czterema kartami wzorcowymi i przykładową kartą bodźcową w elektronicznej wersji zadania WCST*



### Analiza materiału

Raport generowany przez program PEBL dostarcza licznych wskaźników, informujących o poziomie wykonania zadania przez osobę badaną. Wskaźnikami, które najczęściej traktowane są, jako odpowiednie do oceny poziomu elastyczności poznawczej należą „odpowiedzi perseweracyjne” (*perseverative responses*) oraz „błędy perseweracyjne” (*perseverative* 149 Program nie pozwala na wyświetlanie znaków diakrytycznych stosowanych w alfabecie polskim.

errors) (Miles i in., 2021), które raportowane są w postaci liczby odpowiedzi danego typu lub ich odsetka (wskaźnik procentowy). Na potrzeby niniejszego badania, jako wskaźnik elastyczności poznawczej wybrano „procent błędów perseweracyjnych”. Wskaźnik ten brany był również pod uwagę w analizie czynnikowej Miyake i in. (2000)<sup>150</sup>, co pozwala przyjąć, że w złożonym zadaniu poznawczym jakim jest WCST, za procent błędów perseweracyjnych w największym stopniu odpowiada właśnie funkcja giętkości poznawczej (inaczej przełączanie, ang. *shifting*).

Udzielanie odpowiedzi losowych lub tendencyjne wybieranie jednej karty wzorcowej w odpowiedzi na wszystkie karty bodźcowe generuje poprawność odpowiedzi na poziomie 25%. Z tego względu, jako kryterium włączenia wyników do dalszych analiz statystycznych przyjęto próg 30% poprawnych odpowiedzi. Zgodnie z powyższym do weryfikacji hipotez badawczych użyto wyniku „procent błędów perseweracyjnych” uzyskanego przez osoby badane, których odpowiedzi przekroczyły próg poprawności na poziomie 30%. Zadanego kryterium nie spełniło czworo dzieci oraz jedna osoba z grupy młodzieży. Wypełnili je natomiast wszyscy dorośli uczestnicy badania.

Tabela 13 przedstawia średnią poprawność wykonania zadania oraz średni procent błędów perseweracyjnych w poszczególnych grupach wiekowych, obliczone na podstawie wyników uzyskanych przez osoby, których odpowiedzi zakwalifikowano do dalszej analizy statystycznej według podanego powyżej kryterium.

**Tabela 13**

*Średnie i odchylenia standardowe dla procenta poprawnych odpowiedzi oraz procenta błędów perseweracyjnych w teście WCST dla trzech grup wiekowych.*

Grupa wiekowa	Poprawność	Błędy perseweracyjne
	<i>M (SD)</i>	<i>M (SD)</i>
Dzieci ( <i>n</i> =46)	57,8% (11,6)	20,9% (7,3)
Młodzież ( <i>n</i> =49)	63,2% (9,4)	17,7% (5,6)
Dorośli ( <i>n</i> =50)	72,9% (6,3)	13,3% (4,4)

*Adnotacja.* *M* – średnia; *SD* – odchylenie standardowe; *n* – liczebność próby

<sup>150</sup> Należy odnotować, że Miyake i in. (2000) raportowali „liczbę błędów perseweracyjnych” (*number of perseverative errors*) u osób badanych, a nie odsetek lub procent tych błędów. Korzystanie ze wskaźników procentowych ma tę zaletę, że pozwala porównać wyniki z badania testem sortowania kart w wersji standardowej, jak i skróconej (Miles i in., 2021)

#### 5.4.2.2.2 Zadanie n-wstecz

Zadanie n-wstecz (*n-back*) należy do procedur badawczych, które wykorzystywane są do oceny odświeżania pamięci roboczej (*updating*), gdyż wymaga od osób badanych „przechowywania informacji i jednoczesnego monitorowania pojawiających się bodźców” (Piotrowski, Stettner i in., 2009, s. 58) oraz uaktualniania przechowywanej informacji wraz z napływem nowych bodźców. Procedura polega na tym, że osobie badanej przedstawia się kolejno bodźce (np. litery, figury geometryczne), a jej zadaniem jest stwierdzić, czy aktualnie prezentowany bodziec jest taki sam (lub inny), jak bodziec występujący tuż przed nim (1-wstecz) albo poprzedzający go o 2 lub 3 pozycje w sekwencji bodźców (2-wstecz lub 3-wstecz). Im większy dystans występuje między bodźcami do porównania, tym bardziej obciążona jest pamięć robocza i obserwowana jest większa aktywność kory przedczołowej (Braver i in., 1997).

#### Opis zadania

Uczestnikom badania na ekranie komputera wyświetlane są serie czarnych liter na białym tle. W każdej kolejnej próbie osoba badana ma za zadanie ocenić, czy wyświetlana aktualnie litera jest taka sama jak:

- a) litera bezpośrednio ją poprzedzająca – w warunku 1-wstecz;
- b) przedostatnia litera z serii (litera przed literą bezpośrednio poprzedzającą wyświetlaną literę bodźcową) – w warunku 2-wstecz.

Najpierw wyświetlane są próby w warunku 1-wstecz, a następnie w warunku 2-wstecz. W każdym z warunków wyświetla się 3 bloki liter, a każdy z bloków składa się z serii 24 liter. Łącznie zadanie składa się ze 144 prób przebiegających w 6 seriach (24 próby x 3 bloki x 2 warunki). Dodatkowo przed pierwszymi blokami warunków 1-wstecz oraz 2-wstecz przedstawiana była instrukcja oraz seria treningowa, pozwalająca na zapoznanie się z procedurą.

Ocena zgodności litery bodźcowej z materiałem przechowywanym w pamięci dokonywana jest przez osoby badane poprzez wciśnięcie odpowiedniego przycisku SHIFT (podobnie jak w zadaniu *Monitorowania sylab* odpowiedź „TAK” wymaga wciśnięcia prawego klawisza SHIFT, a odpowiedź „NIE” wymaga wciśnięcia lewego klawisza SHIFT). Na dokonanie oceny osoba badana ma 2 sekundy, przy czym każda litera bodźcowa pozostaje na ekranie przez 700 ms, a przed wyświetleniem kolejnej litery ekran pozostaje pusty przez 1300 ms (interwał ten rozpoczyna każdą kolejną próbę).

W każdej serii 24 liter, 1/3 z nich wymaga odpowiedzi twierdzącej. Porządek wyświetlania liter jest taki sam dla każdej osoby badanej i został ustalony na podstawie losowo wygenerowanej kolejności liter, dla każdego bloku osobno.

Całość procedury została zaprogramowana przez autora rozprawy w programie DMDX (Forester i Forester, 2003).

## **Analiza danych**

Raport wynikowy generowany przez DMDX dla każdej osoby badanej zawiera informację o poprawności odpowiedzi oraz czasie reakcji w każdej próbie. Jak podają Piotrowski, Stettner i in. (2009, s. 60) „podstawową zmienną zależną w zadaniu n-wstecz jest poprawność”, dlatego za wskaźnik poziomu odświeżania pamięci roboczej (*updating*) przyjęto procent poprawnych odpowiedzi.

W pierwszym kroku analizy danych obliczono średnią poprawność odpowiedzi każdej osoby badanej osobno dla warunku 1-wstecz i 2-wstecz (z podziałem na odpowiedzi twierdzące i przeczące, jak i łącznie). Warunek 2-wstecz okazał się bardzo wymagający, szczególnie dla uczestników w grupie dzieci. Średnia poprawności odpowiedzi w tej grupie wiekowej wyniosła  $M=52,6\%$  ( $SD = 13,9$ ), a najmniejsza zarejestrowana wartość, to 23% poprawnych odpowiedzi. W grupie młodzieży było to odpowiednio  $M=63\%$  ( $SD=13,6$ ) oraz min. 30%, a w grupie dorosłych  $M=76\%$  ( $SD = 11,4$ ) i min.38%.

Aby wyeliminować odpowiedzi udzielane w sposób tendencyjny (np. naciskanie wyłącznie klawisza sygnalizującego odpowiedź negatywną dałoby średnio 16/24 poprawne odpowiedzi w każdym bloku) lub zupełnie przypadkowy, przyjęto następujące kryterium kwalifikacji odpowiedzi danej osoby do analizy: Do analizy statystycznej włączane były odpowiedzi osoby, która uzyskała próg poprawności odpowiedzi na bodźce docelowe zgodne (odpowiedź poprawna: TAK) na poziomie min. 50% oraz próg poprawności odpowiedzi na bodźce niezgodne (odpowiedź poprawna: NIE) także na poziomie min. 50%. Kryterium to zostało spełnione w zadowalającym stopniu dla warunku 1-wstecz we wszystkich grupach wiekowych (odpowiedzi 90% badanych osób osiągnęły założony próg poprawności), natomiast w warunku 2-wstecz próg ten osiągnęło jedynie 53% badanych<sup>151</sup>. Z tego względu jako wskaźnik funkcji odświeżania pamięci roboczej w dalszych analizach statystycznych na potrzeby weryfikacji hipotez wykorzystany został wynik uzyskany przez osoby badane we wszystkich grupach wiekowych w warunku 1-wstecz. Danych z warunku 2-wstecz nie uwzględniono w analizie. Szczegółowe zestawienie średnich i odchyłeń standardowych dla

---

151 W większości zadane kryteria włączenia spełniły odpowiedzi osób badanych z grupy dorosłych (ponad 90%), natomiast z grup dzieci i młodzieży osiągnęło je odpowiednio 22% oraz 45% uczestników.

procenta poprawnych odpowiedzi w warunkach 1- i 2-wstecz obliczone po wyłączeniu odpowiedzi osób, które nie spełniły zadanego kryterium przedstawia tabela 14.

**Tabela 14**

*Średnie i odchylenia standardowe dla procenta poprawnych odpowiedzi w zadaniu 1- i 2-wstecz*

Grupa wiekowa	Procent poprawnych odpowiedzi 1-wstecz		Procent poprawnych odpowiedzi 2-wstecz	
	<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>
Dzieci	39	79,1 (8,9)	11	65,7 (10,2)
Młodzież	46	89 (5,8)	22	73,8 (6,5)
Dorośli	50	94,6 (3,8)	46	78,2 (1,3)

*Adnotacja:* Uwzględniono wyłącznie wyniki osób, które spełniły kryterium poprawności odpowiedzi twierdzących i przeczących na poziomie minimum 50%; *M* – średnia; *SD* – odchylenie standardowe; *n* – liczebność próby

#### 5.4.2.2.3 Zadanie z flankerami

Zadanie z flankerami (*flanker task*) zostało opracowane przez Eriksen i Eriksena (1974). Polega ono na identyfikowaniu bodźców centralnych wyświetlanych na ekranie i ignorowaniu otaczającego je szumu (bodźce flankery). Reakcja na wyświetlane bodźce ma charakter reakcji wyboru, gdyż w zależności od typu bodźca centralnego należy nacisnąć jedną z dwóch dźwigni (klawiszy). Manipulacja rodzajem szumu wpływa na poprawność oraz czas reakcji, co czyni to zadanie uznanym narzędziem pomiaru uwagi selektywnej i kontroli interferencji (Diamond, 2013; Nigg, 2000; Ridderinkhof i in., 2021).

W oryginalnym zadaniu (Eriksen i Eriksen, 1974) jako bodźce centralne oraz flankery wykorzystano wielkie litery alfabetu: „H” i „K” oraz „S” i „C”. Litery mogły pojawiać się pojedynczo na ekranie (bodziec docelowy bez szumu) lub w towarzystwie tych samych albo innych liter, np. bodziec centralny otoczony przez identyczne bodźce „H H H H H H H” (warunek zgodności) albo bodziec centralny otoczony przez bodźce wymagające przeciwstawnej reakcji „K K K S K K K”. Współcześnie najczęściej występujący wariant (Ridderinkhof i in., 2021) polega na identyfikowaniu kierunku figury centralnej w formie strzałki, która prezentowana jest w dwóch rodzajach orientacji względem figur sąsiadujących: orientacja zgodna (<<<<<<) i niezgodna (<<><<). Kontrolnie stosowany jest także warunek

neutralny, w którym centralne strzałki występują samodzielnie (bez flankerów) lub otoczone są przez poziome linie (bez grotów).

### **Opis zadania**

W projekcie skorzystano z procedury zadania Flanker Task, wersja 1.0 przygotowanego w programie PEBL wersja 0.14 (Mueller i Piper, 2014). W zadaniu występują 3 warunki definiowane przez 3 rodzaje układu bodźców:

- 1) warunek neutralny – strzałka centralna otoczona z każdej strony przez dwie poziome kreski (bez grotów);
- 2) warunek zgodności – strzałka centralna otoczona z każdej strony przez dwie strzałki wskazujące ten sam kierunek;
- 3) warunek niezgodności – strzałka centralna otoczona z każdej strony przez dwie strzałki wskazujące przeciwny kierunek.

W każdym warunku występuje po 40 prób prezentowanych w losowej kolejności. Czas między próbami wynosi 1 sekundę, a punkt fiksacji wzroku, występujący bezpośrednio przed bodźcami docelowymi prezentowany jest przez 400 milisekund. Długość bodźców docelowych (strzałek) wynosi 30 pikseli, a odległość między nimi jest równa 3 pikselom (przy rozdzielczości ekranu 800x600 pikseli). Odpowiedzi na bodźce docelowe w postaci wskazania kierunku figury (strzałki) centralnej udzielane są za pomocą klawiatury: prawy klawisz SHIFT oznacza kierunek prawy, a lewy klawisz SHIFT oznacza wybór kierunku lewego. Główna część zadania poprzedzona jest blokiem testowym, składającym się z kilkunastu prób (minimum 12 prób i wymagane przynajmniej 8 odpowiedzi poprawnych, aby zakończyć blok testowy). Po każdej próbie testowej osoba badana otrzymuje informację zwrotną o poprawności odpowiedzi oraz o czasie reakcji. Po zapoznaniu się z procedurą i wyrażeniu gotowości do wzięcia udziału w zadaniu, każdy z uczestników przechodzi do części głównej, w której nie jest już udzielana informacja zwrotna o poprawności i szybkości odpowiedzi.

### **Analiza danych**

Program PEBL rejestrował liczbę błędów, poprawność odpowiedzi w warunku zgodności i niezgodności oraz czas reakcji w każdym z warunków, a także koszt konfliktu w postaci różnicy między czasem reakcji w warunku zgodności, a czasem reakcji w warunku niezgodności. Punkt czasowy, którego przekroczenie traktowane było, jak udzielenie błędnej

odpowiedzi wynosił 800 ms. Do tego momentu program rejestrował poprawność odpowiedzi oraz czas reakcji, a po przekroczeniu tego progu, odpowiedź automatycznie klasyfikowana była jako błędna. Z tego względu zamiast analizować czasy reakcji odpowiedzi poprawnych, do obliczenia poziomu funkcji hamowania postanowiono wykorzystać informacje dokładności (*accuracy*) odpowiedzi, czyli odsetku odpowiedzi poprawnych. Parametr ten zawiera informację zarówno o poprawności odpowiedzi, a także o czasie reakcji w tym sensie, że wolniejsze reagowanie generowało większy poziom błędu.

Do dalszej analizy na potrzeby statystycznej weryfikacji hipotez włączono wyniki osób, które w warunku kongruencji (zgodności) poprawnie wskazały ponad 50% odpowiedzi oraz średnio w całym zadaniu osiągnęły poprawność na poziomie przekraczającym 50%. Kryterium to przyjęto, aby wykluczyć wyniki uzyskane przez tendencyjne lub całkowicie przypadkowe udzielanie odpowiedzi. W związku z powyższym z dalszej analizy wyłączono wyniki czworga uczestników z pierwszej grupy wiekowej (dzieci). Wszyscy badani z pozostałych dwóch grup wiekowych spełnili zdane kryterium (por. tabela 15).

Jako wskaźnik poziomu kontroli hamowania wykorzystano koszt konfliktu, czyli różnicę między poprawnością odpowiedzi w warunku zgodności a poprawnością odpowiedzi w warunku niezgodności. Dla sprawdzenia, czy poprawność odpowiedzi w warunku zgodności jest istotnie wyższa niż w warunku niezgodności przeprowadzono test znaków rangowych Wilcoxon dla prób zależnych osobno dla każdej z grup wiekowych. Wyniki testu Wilcoxon wykazały istotne różnice między poprawnością odpowiedzi w warunku zgodności i warunku niezgodności w każdej z grup wiekowych (dzieci:  $Z=-5,83$ ;  $p<0,001$ ; młodzież:  $Z=-5,42$ ;  $p<0,001$ ; dorośli:  $Z=-5,86$ ;  $p<0,001$ ; ). Wynik ten potwierdza użyteczność różnicy poprawności odpowiedzi między warunkami zgodności i niezgodności (koszt reakcji), jako wskaźnika poziomu hamowania.

### Tabela 15

*Statystyki opisowe dla średnich poprawności reakcji w zadaniu z flankerami*

Grupa wiekowa	Warunek zgodności	Warunek niezgodności
	<i>M (SD)</i>	<i>M (SD)</i>
Dzieci ( $n=46$ )	86,3% (9,7)	60,8% (20,4)
Młodzież ( $n=50$ )	95,7% (6,5)	87,4% (13,6)
Dorośli ( $n=50$ )	99% (2,1)	91% (8)

*Adnotacja.* *M* – średnia; *SD* – odchylenie standardowe;

## 6 Wyniki

### 6.1 Wprowadzenie

W tym rozdziale przedstawione zostały wyniki przeprowadzonych badań, których metodologię opisano w rozdziale poprzednim. Wyniki opisywane są w porządku wyznaczonym przez kolejne hipotezy. W niektórych przypadkach dzięki możliwości zastosowania np. dwuczynnikowej analizy wariancji kilka hipotez testowanych było łącznie. Poza wynikami weryfikującymi hipotezy postawione w rozdziale 5., w niniejszym rozdziale przedstawiono także dodatkowe analizy, które stanowią odpowiedź na pytania i wątpliwości, które pojawiły się w trakcie testowania hipotez. Analizy statystyczne zostały przeprowadzone przy pomocy pakietu IBM SPSS Statistics 28 i 29. W celu ustalenia różnic międzygrupowych preferowanym testem statystycznym była jedno i dwuczynnikowa analiza wariancji ANOVA. Jeśli nie były spełnione warunki umożliwiające wykonanie analizy wariancji w oparciu o średnie, wykorzystywano alternatywne testy nieparametryczne:  $H$  Kruskala-Wallisa,  $W$  Wilcoxona oraz Jonkheere-Terpstry. Do ustalenia związków między badanymi zmiennymi użyty został współczynnik korelacji  $r$  Pearsona. Przyjęty poziom istotności stosowanych testów statystycznych, na podstawie którego podejmowano decyzję o odrzuceniu (lub nie) hipotezy zerowej, to  $\alpha = 0,05$ .

### 6.2 Kontrola selekcji leksykalnej

W celu weryfikacji hipotezy o wzrastającym wraz z wiekiem poziomie kontroli selekcji leksykalnej (w zależności od kontekstu semantycznego) poddano analizie poprawność wykonania zadania *Nazywanie obrazków*. Statystyki opisowe dla uzyskanych wyników przedstawia tabela 16. Według zawartych w niej danych poprawność wykonania zadania była bardzo wysoka i we wszystkich grupach wyniosła dla warunku neutralnego (bez dystrakcji semantycznej) ponad 95% zgodności odpowiedzi z przyjętym kluczem nazywania obrazków. Dane pokazują także stosunkowo wysoką odporność selekcji leksykalnej na interferencję wywołaną dystrakcją semantyczną. Aby odpowiedzieć na pytania badawcze dotyczące różnic w kontroli selekcji leksykalnej ze względu na wiek oraz kontekst semantyczny postanowiono dokonać porównań między grupami oraz warunkami badania.

**Tabela 16***Statystyki opisowe poprawności nazywania obrazków*

Grupa wiekowa	<i>M</i>	<i>SD</i>	<i>Me</i>	<i>SKE</i>	<i>K</i>
<i>Warunek neutralny</i>					
Dzieci ( <i>n</i> = 50)	96,72	3,35	97,73	-1,02	0,47
Młodzież ( <i>n</i> = 50)	99,21	1,7	100	-2,44	6,23
Dorośli ( <i>n</i> = 50)	99,63	0,92	100	-2,14	2,68
<i>Warunek interferencji</i>					
Dzieci ( <i>n</i> = 50)	93,88	6,59	94,74	-1,40	2,15
Młodzież ( <i>n</i> = 50)	98,60	2,18	100	-1,55	1,72
Dorośli ( <i>n</i> = 50)	99,58	0,98	100	-1,91	1,74

*Adnotacja:* Tabela przedstawia wartości procentowe poprawności nazywania obrazków; *n* – liczba obserwacji; *M* – średnia; *SD* – odchylenie standardowe; *Me* – mediana, *SKE* – skośność; *K* – kurtoza.

Rozkłady wyników w każdym przypadku (w każdej grupie w obu warunkach) są wyraźnie lewoskośne, co oznacza że większość wyników koncentruje się powyżej średniej ( $Me > M$ ). Z powodu znacznej asymetrii oraz wysmuklenia rozkładu wyników względem rozkładu normalnego testowanie hipotez 1.1.1.-1.1.3. postanowiono przeprowadzić za pomocą testów nieparametrycznych (opartych na rangowaniu)<sup>152</sup>.

Podstawowym efektem dystrakcji semantycznej w zadaniu nazywania obrazków jest spowolnienie nazywania określane jako efekt interferencji semantycznej. Hipotezy 1.1.4-1.1.6 odnosiły się do różnic w czasie nazywania obrazków wynikających z wieku osób badanych oraz warunku eksperymentalnego (manipulacja kontekstem semantycznym). Tabela 17 przedstawia średnie czasy nazywania obrazków we wszystkich grupach wiekowych dla dwóch warunków badania (warunek neutralny vs. warunek interferencji), a także różnice między średnimi czasu nazywania.

<sup>152</sup> Warunek normalności rozkładu wyników, konieczny do przeprowadzenia analizy wariancji nie został spełniony w żadnej grupie w żadnym z dwóch warunków (uzyskano wyniki testu Shapiro-Wilka istotne statystycznie  $p < 0,001$ ). Nie udało się zbliżyć rozkładów zmiennej do rozkładu normalnego również po przeprowadzeniu transformacji danych opartej o *arcus sinus* (*arcsine transformation*) zalecanej przez Cohena i in. (2003). Z tego względu wyniki podawane są w postaci podstawowej (przed transformacją) i w dalszej kolejności na ich podstawie testowane będą hipotezy badawcze.

**Tabela 17***Statystyki opisowe dla średnich czasu nazywania obrazków*

Grupa wiekowa	<i>M</i>	<i>SD</i>	<i>Me</i>	<i>SKE</i>	<i>K</i>	<i>W</i>	<i>p</i>
<i>Warunek neutralny</i>							
Dzieci ( <i>n</i> = 50)	942,20	111,67	933	0,45	-0,14	0,97	0,34
Młodzież ( <i>n</i> = 50)	695,60	90,11	687	0,27	-0,87	0,96	0,12
Dorośli ( <i>n</i> = 50)	601,04	63,13	601	-0,06	0,69	0,97	0,31
<i>Warunek interferencji</i>							
Dzieci ( <i>n</i> =50)	1324,08	172,96	1308	0,01	-,075	0,96	0,10
Młodzież ( <i>n</i> = 50)	917,80	131,93	920	0,35	-0,32	0,98	0,47
Dorośli ( <i>n</i> = 50)	739,86	107,52	721,5	0,41	0,13	0,98	0,72
<i>Różnica średnich (Wi-Wn)</i>							
Dzieci ( <i>n</i> = 50)	381,88	130,17	365,5	0,31	-0,43	0,98	0,51
Młodzież ( <i>n</i> = 50)	222,20	92,36	221,5	0,21	-0,61	0,98	0,50
Dorośli ( <i>n</i> = 50)	138,82	73,51	152	0,15	-0,50	0,97	0,33

*Adnotacja.* Czas nazywania podawany jest w milisekundach od wyświetlenia obrazka do rozpoczęcia fonacji; *n* – liczba obserwacji; *M* – średnia; *SD* – odchylenie standardowe; *Me* – mediana, *SKE* – skośność; *K* – kurtoza; *W* – wynik testu Shapiro-Wilka; *p* – wartość *p* dla testu Shapiro-Wilka

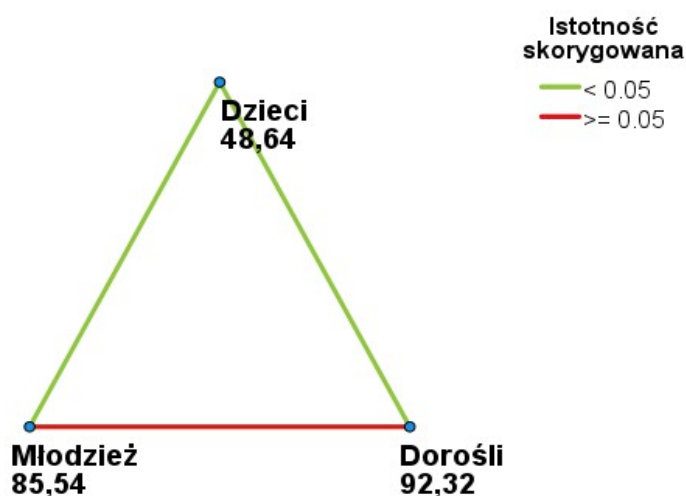
Rozkład średnich czasów nazywania obrazków nie odbiega znacząco od rozkładu normalnego: bezwzględna wartość skośności w żadnej z grup nie przekracza 0,5, a bezwzględna wartość kurtyczności jest poniżej wartości 1 (George i Mallery, 2019; por. Niewiarowski i in., 2013; Wiktorowicz i in., 2020). Zbliżony rozkład zmiennej zależnej potwierdza także wynik testu Shapiro-Wilka, który nie wykazuje istotności w żadnym z sześciu testowanych pomiarów ( $p > 0,05$ ), co nie pozwala odrzucić hipotezy zerowej o zgodności rozkładów z rozkładem normalnym. Spełnienie założenia o normalności rozkładu analizowanych zmiennych, umożliwiło zastosowanie testów parametrycznych w porównaniu średnich czasów reakcji do zweryfikowania hipotez 1.1.4.-1.1.6.

### 6.2.1 Test hipotezy 1.1.1. Skuteczność selekcji leksykalnej wzrasta wraz z wiekiem.

Na początku, aby odpowiedzieć na pytanie, czy skuteczność selekcji leksykalnej zmienia się (wzrasta) wraz z wiekiem przeprowadzono test Kruskala-Wallisa dla prób niezależnych, porównując międzygrupowo poprawność odpowiedzi w warunkach neutralnych. Uzyskany wynik:  $H(2, N = 150) = 41,36$ ;  $p < 0,001$  pozwala na odrzucenie hipotezy zerowej o braku istotnych różnic w rozkładach poprawności odpowiedzi między grupami. Wielkość efektu obliczona na podstawie statystyki  $H$  (Tomczak i Tomczak, 2014) wyniosła  $\eta^2_H = 0,268$ . Porównania *post hoc* pomiędzy poszczególnymi grupami w parach przy poziomie istotności skorygowanej metodą Bonferroni dla wielu porównań wykazały, że poziom poprawności wykonania zadania różni się w sposób istotny ( $p < 0,05$ ) między grupą młodzieży i dzieci (różnica średnich rang wyniosła 36,9) oraz grupą dorosłych i dzieci (różnica średnich rang wyniosła 43,68). Pomiedzy grupą dorosłych i młodzieży różnica była nieistotna statystycznie (różnica pomiędzy średnimi rang wyniosła 6,78). Poniżej na rysunku 16 przedstawiono wykres obrazujący uzyskane wyniki porównań międzygrupowych. Na ich podstawie można stwierdzić, że skuteczność selekcji leksykalnej mierzona poprawnością nazywania obrazków wzrasta wraz z wiekiem, osiągając swoje maksimum już w adolescencji. Oznacza to, że **hipoteza 1.1.1. została częściowo potwierdzona**, gdyż „efekt sufitu” osiągnięty dla tego zadania przez badaną młodzież nie pozwala na stwierdzenie różnic rozwojowych między grupą adolescentów i młodymi dorosłymi.

#### Rysunek 16

Wyniki porównań poprawności nazywania obrazków w warunkach neutralnych między grupami wiekowymi



Uwaga: Liczby przy węzłach reprezentują wartości średnich rang wszystkich obserwacji w poszczególnych próbach

### **6.2.2 Test hipotezy 1.1.2. Pod wpływem zewnętrznej dystrakcji semantycznej skuteczność selekcji leksykalnej obniża się u przedstawicieli wszystkich badanych grup wiekowych.**

Aby sprawdzić, czy skuteczność selekcji leksykalnej pod wpływem zewnętrznej dystrakcji semantycznej obniża się u osób w okresie późnego dzieciństwa, adolescencji i dorosłości postanowiono porównać poprawność odpowiedzi w zadaniu nazywania obrazków w warunku neutralnym i warunku interferencji. Ze względu na to, że w grupie dorosłych poprawność odpowiedzi w obu warunkach niemal się nie różni (różnica wynosi 0,05 punktu procentowego, zob. tabela 16), postanowiono przeprowadzić osobne porównania w każdej z trzech grup wiekowych przy wykorzystaniu nieparametrycznego testu Wilcoxon dla prób zależnych.

Analiza testem znaków rangowych Wilcoxon wykazała istotną różnicę w poprawności odpowiedzi pomiędzy warunkami (warunek neutralny vs. warunek interferencji) jedynie w grupie dzieci:  $Z = -3,37$ ;  $p < 0,001$ ;  $r_c = -0,48$ . Różnice występujące pomiędzy warunkami w grupie młodzieży ( $Z = -1,79$ ;  $p = 0,073$ ;  $r_c = -0,25$ ) oraz w grupie dorosłych ( $Z = -0,5$ ;  $p = 0,617$ ;  $r_c = -0,07$ ) okazały się nieistotne statystycznie<sup>153</sup>. Wyniki te w odniesieniu do hipotezy 1.1.2. można interpretować, uznając, że pod wpływem zewnętrznej dystrakcji semantycznej skuteczność selekcji leksykalnej obniża się jedynie u osób z grupy dzieci. Manipulacja kontekstem semantycznym i wprowadzenie dystrakcji semantycznej nie powoduje natomiast istotnego obniżenia się wyników w grupie młodzieży i grupie osób dorosłych. Tym samym można przyjąć, że **hipoteza 1.1.2 również potwierdzona została jedynie częściowo.**

### **6.2.3 Test hipotezy 1.1.3. Poziom odporności selekcji leksykalnej na zewnętrzną dystrakcję semantyczną wzrasta wraz z wiekiem.**

Otrzymane w poprzednich analizach wyniki pokazują, że poszczególne grupy badanych osób różnią się odpornością selekcji leksykalnej na dystrakcję semantyczną. Należy zauważyć, że badani w każdej z grup prezentowali bardzo wysoki poziom odporności selekcji leksykalnej na zewnętrzną dystrakcję semantyczną, dzięki czemu poprawność nazywania w warunku interferencji nie różniła się znacznie od poziomu poprawności w warunku neutralnym, a w grupie dorosłych i młodzieży różnica ta okazała się nieistotna statystycznie. Można więc spodziewać się, że ewentualny wzrost odporności selekcji leksykalnej na

---

<sup>153</sup> W grupie dorosłych na 50 porównań w teście znaków rangowych Wilcoxon, aż w 37 przypadkach wyniki w obu warunkach były identyczne (37 wiązań).

dystrakcję wraz z wiekiem jest nieznaczny, gdyż już stosunkowo wcześniej obserwuje się niską podatność na dystrakcję (w zakresie poprawnego nazywania obrazków). Ponadto, jak wykazano przy weryfikacji hipotezy 1.1.1. skuteczność selekcji leksykalnej w warunku neutralnym nie różni się istotnie u osób badanych z grup młodzieży i dorosłych, co przy obserwacji tak niewielkich różnic między warunkami powoduje, że wzrost poziomu odporności selekcji leksykalnej na dystrakcję wraz z wiekiem może być trudny do uchwycenia w analizowanej próbie. Aby sprawdzić, czy pomimo niewielkich różnic uda się zaobserwować istotny trend rozwojowy przeprowadzono nieparametryczny test Jonckheere-Terpstry, który jest odpowiedni do sprawdzenia trendów w danych porządkowych (Conover, 1999; Field, 2009) i dla porównań z założoną *a priori* tendencją wzrastającą (lub malejącą) wykazuje większą moc statystyczną niż zastosowany przy weryfikacji hipotezy 1.1.1. test Kruskala-Wallisa (Hollander i in., 2014). Otrzymano wynik  $T_{JT} = 2952$ ;  $z = -2,95$ ;  $p = 0,003$ ;  $\tau_b = -0,204$ <sup>154</sup>. Istotny wynik testu oznacza, że zakładany trend rozwojowy znalazł potwierdzenie w danych, a odporność selekcji leksykalnej na dystrakcję zwiększa się wraz z wiekiem (ujemny trend oznacza, że zmniejsza się różnica między poprawnością odpowiedzi w warunku neutralnym w stosunku do warunku interferencji). Efekt ten jest jednak stosunkowo słaby (por. Ferguson, 2009). Dalsze porównania *post hoc* w parach wykazały, że różnice między dziećmi i młodzieżą oraz grupą młodzieży a dorosłymi nie były istotne po korekcie Bonferroniego<sup>155</sup> (odpowiednio:  $p = 0,139$  i  $p = 0,135$ ), a jedyna istotna różnica wystąpiła między grupą dzieci, a grupą dorosłych ( $p = 0,018$ ). Na podstawie otrzymanych wyników można więc wnioskować, że odporność selekcji leksykalnej na dystrakcję semantyczną zwiększa się tylko gdy porównywane są wyniki dzieci i osób dorosłych, a jej wzrost między dzieciństwem i adolescencją oraz między okresem dorastania i wczesnej dorosłości może nie być zauważalny (przynajmniej w warunkach sztucznie wygenerowanej interferencji podczas nazywania obrazków). Tym samym **można uznać hipotezę 1.1.3. za potwierdzoną**, choć jednocześnie zaobserwowany trend rozwojowy jest niewielki, a różnice między poszczególnymi grupami wiekowymi są bardzo subtelne.

---

154 Jako miarę siły efektu zastosowano współczynnik tau b Kendalla, który przyjmuje wartości  $\langle -1, 1 \rangle$ . Według Conovera (1999, s. 325) test obliczający tau Kendalla jest równoważny (*equivalent*) testowi Jonckheere-Terpstry. Podobną procedurę statystyczną, polegającą na wykorzystaniu tau Kendalla jako miary siły efektu dla testu JT zastosowali np. Maycock i in. (2018).

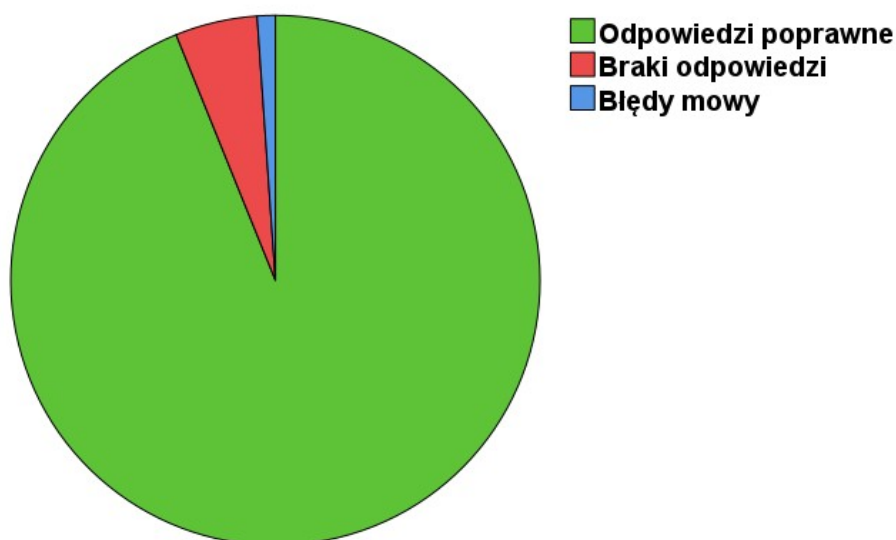
155 Przed korektą istotność osiągała wartość  $p=0,046$  dla porównania wyników w grupie dzieci i młodzieży oraz  $p=0,045$  dla porównań wyników młodzieży oraz dorosłych.

### 6.2.3.1 Dodatkowa analiza: Stosunek błędów mowy i braków odpowiedzi do ogólnej poprawności odpowiedzi

Choć nie było to przedmiotem stawianych hipotez, warto sprawdzić w jaki sposób zewnętrzna dystrakcja semantyczna wpłynęła na obniżenie skuteczności selekcji leksykalnej (u dzieci): czy spowodowała więcej błędów mowy w postaci wypowiedziania słów-dystraktorów (lub innych przejęzyczeń), czy raczej doprowadziła do zablokowania wydobywania właściwej lemmy i nienazwania obrazka (próba zaklasyfikowana jako „brak odpowiedzi”)?

#### Rysunek 17

*Średni procentowy rozkład odpowiedzi udzielanych przez dzieci w warunkach interferencji*



Wykres na rysunku 17 zestawia poprawne odpowiedzi (93,88%) w najmłodszej badanej grupie wiekowej z brakami odpowiedzi (5,02%) oraz z błędami mowy (1,1%). Widać na nim, że braki odpowiedzi stanowią dominującą grupę wśród odpowiedzi niezaklasyfikowanych jako poprawne.

Aby w grupie dzieci porównać między warunkami (warunek neutralny vs. warunek interferencji) rozkłady braków odpowiedzi oraz błędów mowy przeprowadzono dwukrotnie test Wilcozona dla danych zależnych. W przypadku braków odpowiedzi uzyskano istotny statystycznie wynik:  $Z = -4,3$ ;  $p < 0,001$ ;  $r_c = -0,61$ , co oznacza, że w odpowiedziach

udzielanych przez dzieci w warunku interferencji znacząco częściej występował brak podania nazwy obrazka niż w warunku neutralnym. W przypadku porównania odsetka błędów mowy okazało się, że nie ma istotnej różnicy między warunkami w częstotliwości przejęzyczeń popełnianych przez dzieci, a otrzymany wynik nie uzyskał zadanego poziomu istotności:  $Z = -0,93$ ;  $p = 0,36$ ;  $r_c = -0,13$ . Dane te można interpretować następująco: Wykazane w grupie dzieci istotne obniżenie skuteczności selekcji leksykalnej pod wpływem zewnętrznej dystrakcji semantycznej związane było z zablokowaniem tego procesu przez wyraz dystrakcyjny, którego aktywacja uniemożliwiała skuteczne wydobycie docelowej jednostki leksykalnej, jednocześnie nie powodując istotnego wzrostu odsetka przejęzyczeń leksykalnych.

#### **6.2.4 Test hipotez 1.1.4.-1.1.6. dotyczących rozwojowych oraz kontekstualnych uwarunkowań szybkości selekcji leksykalnej**

Testowane w tej sekcji hipotezy zostały sformułowane następująco:

- Hipoteza 1.1.4. Szybkość selekcji leksykalnej wzrasta wraz z wiekiem.
- Hipoteza 1.1.5. Szybkość selekcji leksykalnej zmniejsza się pod wpływem zewnętrznej dystrakcji semantycznej u osób w różnym wieku.
- Hipoteza 1.1.6. Wielkość efektu interferencji semantycznej maleje wraz z wiekiem.

W celu sprawdzenia, czy zgodnie z postawionymi hipotezami wiek różnicuje odpowiedzi osób badanych pod względem czasu efektywnej selekcji leksykalnej w zależności od kontekstu semantycznego (w zadaniu nazywania obrazków), przeprowadzona została dwuczynnikowa analiza wariancji w schemacie mieszanym: 3 (grupa wiekowa) x 2 (poziom dystrakcji semantycznej: warunek neutralny vs warunek interferencji). Czynnikiem mierzonym wewnątrz osób był warunek kontekstu semantycznego, natomiast czynnikiem między osobami była przynależność do jednej z grup wiekowych. Zmienną zależną był czas efektywnej selekcji leksykalnej wyznaczany przez czas nazywania obrazków (od wyświetlenia obrazka do jego artykulacji). Analiza wariancji została przeprowadzona w modelu jednozmiennowym (*univariate repeated measures ANOVA*). Uzyskane zostały następujące istotne efekty:

1. efekt główny przynależności do grupy wiekowej  $F(2, 147) = 227,44$ ;  $p < 0,001$ ;  
 $\eta_p^2 = 0,756$ <sup>156</sup>;

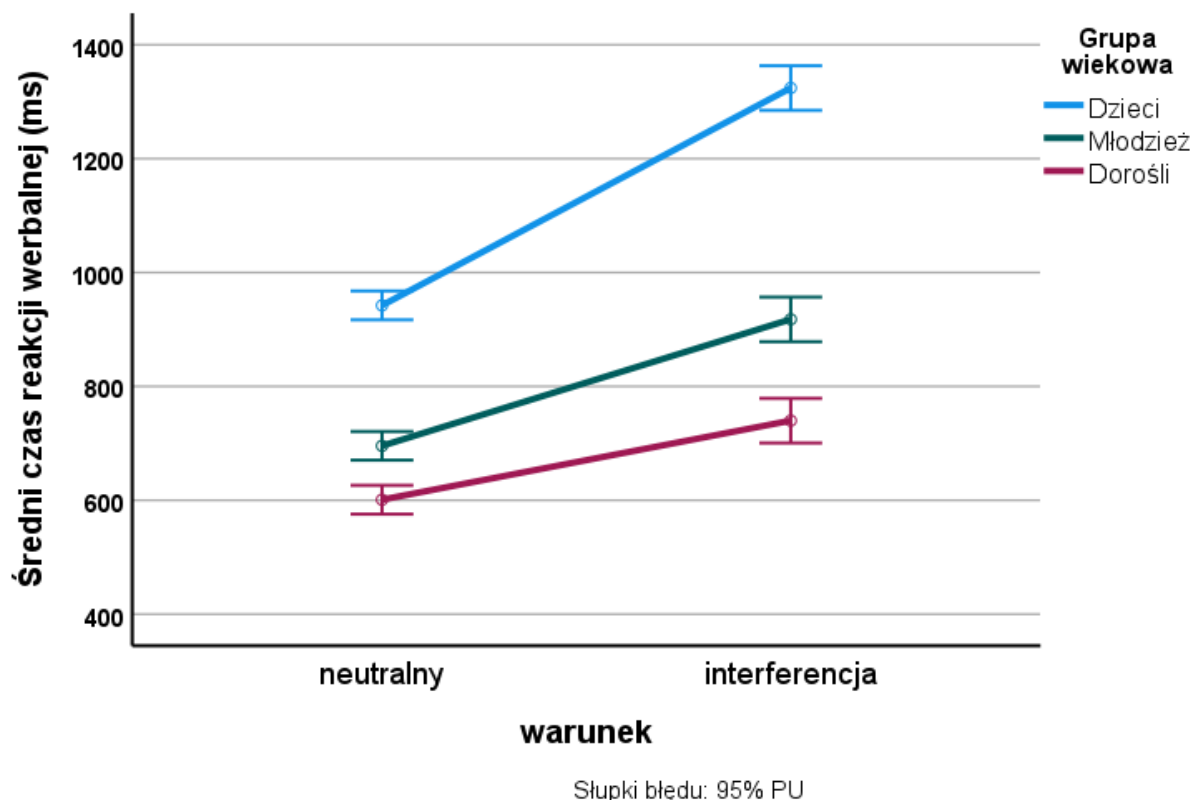
---

<sup>156</sup> Ze względu na niespełnienie założenia o homogeniczności wariancji dla międzyobiektowej zmiennej niezależnej (wartość testu Levene'a równości wariancji błędów była istotna statystycznie) zastosowano poprawkę Welcha. Wartości testu  $F$  oraz  $\eta_p^2$  podane są już w postaci skorygowanej.

2. efekt główny warunku nazywania obrazków (manipulacji kontekstem semantycznym):  
 $F(1,147) = 893,69; p < 0,001; \eta_p^2 = 0,859;$
3. efekt interakcji wieku oraz warunku zadania  $F(2,147) = 74,11; p < 0,001; \eta_p^2 = 0,502.$

## Rysunek 18

Średnie czasów reakcji w zadaniu nazywania obrazków



Ad. 1. Istotny efekt główny przynależności do grupy wiekowej pozwala uznać, że od wieku (przynależności do danej grupy wiekowej) zależy ponad 75% zmienności w szybkości nazywania obrazków (w obu warunkach łącznie). Przeprowadzone porównania *post hoc* za pomocą testu Tamhane'a ujawniły istotne różnice ( $p < 0,001$ ) między wszystkimi grupami.

Aby sprawdzić czy czas efektywnej selekcji leksykalnej (nieobciążonej zewnętrzną dystrakcją semantyczną) jest różny w poszczególnych grupach wiekowych przeprowadzono dodatkowo analizę prostych efektów głównych przynależności do grupy tylko dla warunku neutralnego (pomiar czasu nazywania bez dystrakcji). W tym celu wykonano jednoczynnikową analizę wariancji w planie międzygrupowym oraz porównania *post hoc*. Wartość testu  $F$  z korektą Welcha wyniosła  $F(2, 147) = 175,83; p < 0,001$ . Porównania *post hoc* za pomocą testu Tamhane'a ponownie ujawniły istotne różnice ( $p < 0,001$ ) między

wszystkimi grupami. Wyniki porównań średnich przedstawia tabela 18. Najniższy średni czas nazywania obrazków w warunku neutralnym osiągnął przez osoby dorosłe na poziomie istotnie różnym od wyników w grupie młodzieży i grupie dzieci. Podobnie, średni czas nazywania uzyskiwany w grupie młodzieży był istotnie niższy niż w grupie dzieci, które nazywały obrazki z największym poziomem latencji. Wyniki te pokazują, że czas efektywnej selekcji leksykalnej, mierzony szybkością nazywania obrazków, wykazuje tendencję spadkową wraz z wiekiem, co pozwala uznać, że **hipoteza 1.1.4. została potwierdzona.**

**Tabela 18**

*Porównanie między grupami średnich czasów nazywania obrazków w warunku neutralnym*

Zestawiane grupy wiekowe (A-B)	Różnica średnich	SE	p
Dzieci - młodzież	246,60	20,30	<0,001
Dzieci - dorośli	341,16	18,14	<0,001
Młodzież - dorośli	94,56	15,56	<0,001

*Adnotacja.* SE – błąd standardowy; p – wartość p dla porównań post hoc testem Tamhane’a z poprawką dla porównań wielokrotnych Bonferroniego.

Ad. 2. Średni czas reakcji dla warunku neutralnego wyniósł dla wszystkich grup wiekowych łącznie  $M = 746$  ms ( $SD = 170$ ), a dla warunku z dystrakcją semantyczną  $M = 994$  ms ( $SD = 282$ ). Istotny efekt główny warunku nazywania obrazków (manipulacji kontekstem semantycznym) pozwala uznać, że szybkość nazywania obrazków zależy w ponad 85%<sup>157</sup> od manipulacji eksperymentalnej. Oznacza to, że zgodnie z przewidywaniami czas efektywnej selekcji leksykalnej wydłuża się pod wpływem zewnętrznej dystrakcji semantycznej. Tym samym można uznać, że **hipoteza 1.1.5. została potwierdzona.**

Ad. 3. Aby sprawdzić, czy efekt interferencji semantycznej ma odmienną wielkość u osób w różnych grupach wiekowych, poddano analizie efekt interakcji obu czynników głównych. Z wykresu na rysunku 18 można odczytać, że wraz z wiekiem nie tylko zmniejsza się średni czas nazywania obrazków (w obu warunkach), ale także że w coraz starszych grupach wiekowych wpływ dystrakcji na czas selekcji leksykalnej wykazuje trend malejący. Uzyskany

<sup>157</sup> Należy zauważyć, że miara siły efektu jaką jest  $\eta_p^2$  nie sumuje się do 1 w celu uzyskania całkowitego odsetka wyjaśnionej w modelu wariancji. Nie jest także możliwe porównanie wielkości  $\eta_p^2$  dla różnych źródeł zmienności (Sechrest i Yeaton, 1982). Wynika to z faktu, że wielkość  $\eta_p^2$  zależy od stopnia manipulacji zmiennymi eksperymentalnymi, którą trudno oszacować, zwłaszcza jeśli badanie ma charakter quasi-eksperymentalny i w modelu analizy wariancji znajdują się czynniki nie podlegające manipulacji, np. płeć albo wiek (Richardson, 2011).

istotny efekt interakcji obydwu zmiennych niezależnych (grupa wiekowa oraz poziom dystrakcji semantycznej) oznacza, że w zależności od wieku (przynależności do danej grupy wiekowej) manipulacja kontekstem semantycznym w różnym stopniu wpływała na czas reakcji werbalnej. Aby potwierdzić, czy wraz z wiekiem maleje wielkość efektu interferencji semantycznej (wniosek wyciągnięty na podstawie analizy efektów prostych), przeprowadzono dodatkowo jednoczynnikową analizę wariancji w schemacie międzygrupowym, porównując średnie różnice w czasie reakcji między warunkami (warunek interferencji minus warunek neutralny), będące miarą efektu interferencji. Podstawowe dane na temat rozkładu przedstawiają ostatnie wiersze tabeli 17. Uzyskano następujące rezultaty:  $F(2, 147) = 66,82$ ;  $p < 0,001$ ;  $\eta_p^2 = 0,476$ <sup>158</sup>. Porównania *post hoc* przy użyciu testu Tamhane'a z poprawką na wielokrotne porównania, wykazały istotne różnice ( $p < 0,001$ ) pomiędzy każdą z grup wiekowych. Oznacza to, że średnie różnice między warunkami wraz z wiekiem (porównanie grupy dzieci z grupą młodzieży oraz grupy młodzieży z grupą dorosłych) są coraz mniejsze, przy jednoczesnym zachowaniu istotnego efektu interferencji w obrębie każdej z grup<sup>159</sup>. **Na podstawie uzyskanych wyników można przyjąć, że hipoteza 1.1.6. również została potwierdzona.**

#### **6.2.4.1 Dodatkowa analiza (1): Procentowa wartość opóźnienia nazywania wywołana dystrakcją semantyczną**

Dodatkowo, wykazane wcześniej bezwzględne wartości efektu interferencji semantycznej postanowiono wyrazić procentowo, porównując o ile procent obniża się szybkość nazywania obrazków pod wpływem zewnętrznej dystrakcji semantycznej u uczestników w poszczególnych grupach wiekowych. Uzasadnieniem dodatkowej analizy jest to, że długi czas reakcji u dzieci w nazywaniu bez interferencji i większy niż u młodzieży i dorosłych bezwzględny czas opóźnienia spowodowany dystrakcją nie muszą przekładać się na procentowo większe opóźnienie. Inaczej mówiąc, pomimo wyraźnych różnic w bezwzględnych wartościach czasów reakcji (nazywania), efekt interferencji semantycznej nie musi powodować większych procentowo spadków czasu nazywania w poszczególnych grupach wiekowych. Postanowiono zatem sprawdzić, czy bezwzględnie różne opóźnienie nazywania w trzech grupach wiekowych są proporcjonalne do różnych czasów nazywania osiąganych przez osoby w różnym wieku.

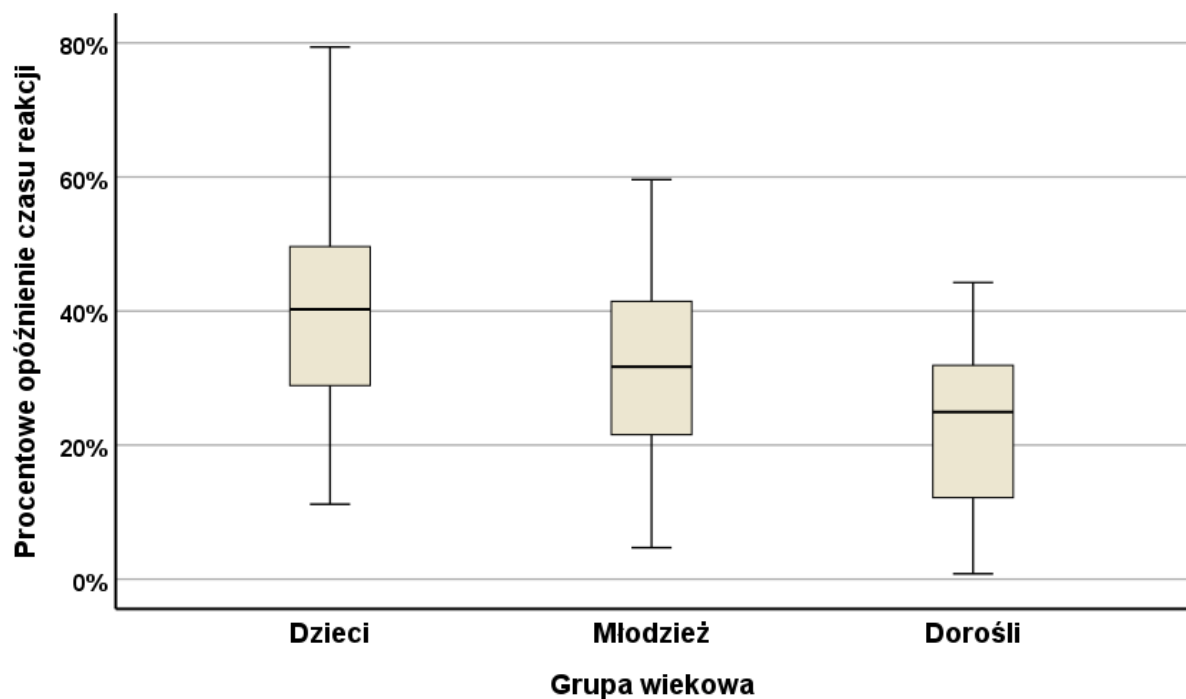
---

158 Ze względu na niespełnienie założenia o homogeniczności wariancji dla międzyobiektowej zmiennej niezależnej (wartość testu Levene'a równości wariancji błędów była istotna statystycznie) zastosowano poprawkę Welcha. Wartości testu  $F$  oraz  $\eta_p^2$  podane są już w postaci skorygowanej.

159 Analiza prostych efektów głównych dla czynnika warunek pokazuje, że różnice w czasach reakcji między warunkami są istotne ( $p < 0,001$ ) w obrębie każdej z grup wiekowych.

## Rysunek 19

Procentowe opóźnienie nazywania w trzech grupach wiekowych wywołane dystrakcją semantyczną



Procentową wartość opóźnienia nazywania z dystrakcją względem warunku neutralnego<sup>160</sup> przedstawia rysunek 19. Wynika z niego, że wraz z wiekiem między kolejnymi okresami rozwojowymi opóźnienie nazywania wywołane interferencją spada o około 9 punktów procentowych z średnio 41% w grupie dzieci, na 32% w grupie adolescentów i 23% w grupie osób dorosłych. Aby sprawdzić, czy różnice te są istotne statystycznie, przeprowadzono jednoczynnikową analizę wariancji w schemacie międzygrupowym uzyskując następujące wyniki:  $F(2, 147) = 22,87$ ;  $p < 0,001$ ;  $\eta_p^2 = 0,237$ . Pozwala to odrzucić hipotezę zerową o braku różnic między grupami w wielkości procentowego opóźnienia wywołanego dystrakcją semantyczną. Porównania *post hoc* przy użyciu testu Scheffègo wykazały istotne różnice ( $p < 0,05$ ) pomiędzy każdą z grup wiekowych. Oznacza, to że zastosowana manipulacja kontekstem semantycznym nie tylko powoduje większy efekt interferencji (wyrażony bezwzględnie dłuższym czasem reakcji w nazywaniu obrazków) w

<sup>160</sup> Procentowe opóźnienie nazywania wywołane dystrakcją nie jest powszechnie używanym wskaźnikiem interferencji, stąd nie jest raportowane w pracach empirycznych wykorzystujących paradygmat *picture-word interference paradigm*.

młodszych grupach wiekowych względem grup starszych, ale także przekłada się na większe procentowe opóźnienie w nazywaniu.

#### **6.2.4.2 Dodatkowa analiza (2): Dystrybucja pauz wypełnionych w odpowiedziach uczestników badania**

Przedstawione do tej pory wyniki są zgodne z ogólną hipotezą o wzrastającym wraz z wiekiem poziomie kontroli selekcji leksykalnej (hipoteza 1.1). Efekt interferencji semantycznej przejawiający się w opóźnieniu nazywania pod wpływem dystrakcji jest wyraźnie większy u młodszej części badanej próby i maleje wraz z wiekiem. W przypadku grupy dzieci dystrakcja semantyczna ma na tyle silny wpływ na selekcję leksykalną, że istotnie przyczynia się do obniżenia poprawności odpowiedzi, zwiększając odsetek braków odpowiedzi w warunku interferencji w porównaniu do warunku neutralnego. Jednocześnie, pomimo pogorszenia poprawności odpowiedzi pod wpływem zewnętrznej dystrakcji semantycznej, skuteczność selekcji leksykalnej pozostaje u dzieci nadal na bardzo wysokim poziomie. Może to świadczyć o tym, że mechanizmy kontrolne na poziomie selekcji leksykalnej działają według zasady preferencji poprawności kosztem płynności mówienia. Tendencja ta powinna mieć zatem odzwierciedlenie także w dystrybucji pauz wypełnionych, poprzedzających wygenerowanie właściwej nazwy obrazka. Można spodziewać się, że w warunku interferencji – przynajmniej w grupie dzieci – procent odpowiedzi z pauzą wypełnioną będzie istotnie wyższy niż w warunku neutralnym, co odzwierciedlałoby zachodzące przedartykulacyjne redagowanie (*prearticulatory editing*) nazwy obrazka i odraczenie jej wypowiedzenia do czasu przewyciężenia interferencji. Aby to sprawdzić przeprowadzono trzy odrębne testy *W* Wilcoxona dla prób zależnych, porównując częstość pauz wypełnionych między warunkami w poszczególnych grupach wiekowych. Tabela 19 przedstawia szczegółowe porównania. Wynika z nich, że tylko w grupie dorosłych różnica między warunkami nie jest istotna statystycznie i niski odsetek pauz wypełnionych jest podobny w obu warunkach zadania nazywania obrazków. W przypadku grup dzieci i młodzieży potwierdziło się przypuszczenie, że dystrakcja semantyczna wpływa na zwiększenie odsetka pauz wypełnionych w warunku interferencji względem warunku neutralnego.

**Tabela 19.**

Wyniki porównań testem Wilcoxon dla dystrybucji pauz wypełnionych między warunkami zadania Nazywanie obrazków w trzech badanych grupach wiekowych

Grupa wiekowa	Warunek	Warunek	Z	p	r <sub>c</sub>
	neutralny	interferencji			
	M (SD)	M (SD)			
Dzieci (n = 50)	1,4 (2,9)	3,9 (7,8)	2,40*	0,016	0,34
Młodzież (n = 50)	2 (5,7)	3,8 (6,5)	2,39*	0,017	0,34
Dorośli (n = 50)	0,9 (3,4)	1,3 (3,9)	0,77	0,44	0,11

Adnotacja: średnia oraz odchylenie standardowe dotyczą procentu pauz wypełnionych w grupach; n – liczebność próby; M – średnia; SD – odchylenie standardowe; R – rangi; Z – wartość testu Wilcoxon; p – wartość p dla testu Wilcoxon; \*p < 0,05; r<sub>c</sub> – siła efektu

### 6.3 Efektywność automonitorowania wewnętrznego

Automonitorowanie procesu mówienia badane było w zadaniu monitorowania sylab (część 2 zadania: monitorowanie wewnętrzne). Poniżej w tabeli 20 zestawiono dane na temat rozkładu poprawnych odpowiedzi (zmienna „efektywność monitorowania wewnętrznego”). Ze względu na znaczne odstępstwo od normalności rozkładu nie było możliwe przeprowadzenie analizy przy pomocy testów parametrycznych na surowych danych. Aby przybliżyć je do rozkładu normalnego przeprowadzono transformację opartą o *arcus sinus* (*arcsine transformation*), która polega na zastosowaniu odwrotności funkcji sinusowej do pierwiastka kwadratowego wartości zmiennej, a następnie pomnożeniu wyniku przez dwa<sup>161</sup> (Cohen i in., 2003). W ten sposób wygenerowano nową zmienną, o rozkładzie bardziej symetrycznym niż dane wyjściowe, która w kolejnym kroku poddana została testom statystycznym. Podstawowe informacje o rozkładzie poprawności wyników w całym zadaniu (oba warunki łącznie) po dokonanym przekształceniu przedstawia tabela 21.

161 Zastosowane przekształcenie oparto na funkcji *arcus sinus* zgodnie ze wzorem  $A = 2 * \arcsin(\sqrt{P})$ , w którym przekształcona zmienna jest równa „dwukrotnej wartości kąta (mierzonego w radianach), którego sinus trygonometryczny jest równy pierwiastkowi kwadratowemu proporcji ulegającej przekształceniu” (Cohen i in., 2003, s. 241).

**Tabela 20***Statystyki opisowe dla poprawności odpowiedzi w zadaniu Monitorowanie sylab*

Grupa wiekowa	<i>M</i>	<i>SD</i>	<i>Me</i>	<i>SKE</i>	<i>K</i>
<i>Warunek zgodności</i>					
Dzieci ( <i>n</i> = 50)	77,40	16,04	82,14	-0,91	0,53
Młodzież ( <i>n</i> = 49)	90,23	6,64	89,29	-0,38	-0,54
Dorośli ( <i>n</i> = 50)	94,92	5,76	96,43	-2,07	6,81
<i>Warunek niezgodności</i>					
Dzieci ( <i>n</i> = 50)	80,14	14,97	82,74	-1,20	1,33
Młodzież ( <i>n</i> = 49)	90,13	8,57	92,59	-2,43	9,26
Dorośli ( <i>n</i> = 50)	97,32	3,71	100	-1,48	2,18

*Adnotacja.* W tabeli podane zostały wartości procentowe poprawnych odpowiedzi; *n* – liczba obserwacji; *M* – średnia; *SD* – odchylenie standardowe; *Me* – mediana, *SKE* – skośność; *K* – kurtoza; *W* – wynik testu Shapiro-Wilka; *p* – wartość *p* dla testu Shapiro-Wilka

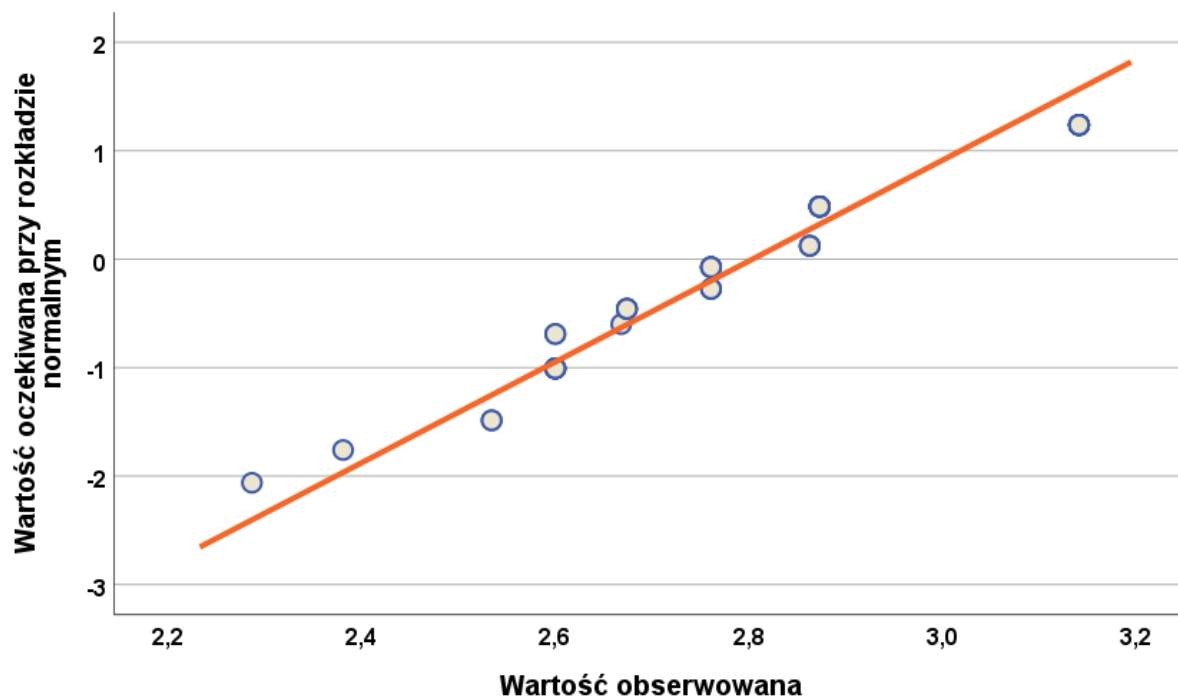
**Tabela 21***Statystyki opisowe poprawności odpowiedzi w zadaniu Monitorowanie sylab po transformacji arcsin w trzech grupach wiekowych*

Grupa wiekowa	<i>M</i>	<i>SD</i>	<i>Me</i>	<i>SKE</i>	<i>K</i>	<i>W</i>	<i>p</i>
Dzieci ( <i>n</i> = 50)	2,23	0,35	2,25	-0,24	0,91	0,98	0,512
Młodzież ( <i>n</i> = 49)	2,54	0,22	2,53	-0,27	1,84	0,96	0,112
Dorośli ( <i>n</i> = 50)	2,8	0,22	2,76	0,1	-0,38	0,92	0,002

*Adnotacja.* *n* – liczba obserwacji; *M* – średnia; *SD* – odchylenie standardowe; *Me* – mediana, *SKE* – skośność; *K* – kurtoza; *W* – wynik testu Shapiro-Wilka; *p* – wartość *p* dla testu Shapiro-Wilka

## Rysunek 20

Wykres normalności K-K z trendem dla rozkładu poprawności odpowiedzi po transformacji arcsin w grupie osób dorosłych



Dzięki zastosowanej transformacji rozkłady wyników przyjęły kształt zbliżony do rozkładu normalnego. Test Shapiro-Wilka przeprowadzony na wynikach z grup dzieci oraz młodzieży okazał się nieistotny statystycznie ( $p > 0,05$ ), co pozwala przyjąć, że rozkłady wyników w tych grupach nie odbiegają w sposób istotny od rozkładu normalnego. Istotny wynik testu Shapiro-Wilka w odniesieniu do próby osób dorosłych nakazuje zachować ostrożność przy interpretacji wyników w oparciu o średnią z tej próby. Jednocześnie wykres K-K (kwantyl-kwantyl) z trendem na rysunku 20 przedstawia akceptowany stopień odstępstwa rozkładu zmiennej zależnej w grupie osób dorosłych od rozkładu normalnego (por. Wiktorowicz i in., 2020, s. 75-76).

Oprócz zbierania danych o poprawności odpowiedzi w zadaniu *Monitorowanie sylab* rejestrowano także czas reakcji, służący za wskaźnik zmiennej „czas efektywnego monitorowania wewnętrznego”. W poniższych tabelach zamieszczono informacje na temat

czasów reakcji<sup>162</sup> w warunkach zgodności (tabela 22) oraz niezgodności (tabela 23) z uwzględnieniem struktury wyrazu (pozycja wykrywanej sylaby w warunku zgodności oraz długość wyrazu w warunku niezgodności). Dla przejrzystości porównania w tabelach nie podano wartości mediany, skośności, ani kurtozy (znaleźć je można w Załączniku G).

W żadnym z warunków ani w żadnej z grup skośność nie przekroczyła wartości bezwzględnej  $|1|$ , a kurtoza nie przekroczyła wartości  $|3|$ . Analiza wyników przy pomocy testu Shapiro-Wilka wykazała, że w większości przypadków rozkłady nie odbiegają w sposób istotny od rozkładu normalnego, a jedynie w warunku zgodności dla rozkładu wyników z próby osób dorosłych dla pomiarów pierwszej (nagłos) i trzeciej (wygłos) sylaby uzyskano istotnie statystyczny ( $p < 0,05$ ) wynik testu Shapiro-Wilka.

**Tabela 22**

*Średnie i odchylenia standardowe dla średnich czasów reakcji w warunku zgodności*

	Dzieci ( $n = 38$ )	Młodzież ( $n = 48$ )	Dorośli ( $n = 50$ )
Pozycja sylaby	$M (SD)$	$M (SD)$	$M (SD)$
Nagłos	1060 (180)	844 (133)	692 (122)
Śródgłos	1373 (222)	1009 (200)	842 (158)
Wygłos	1331 (215)	1001 (161)	821 (150)

*Adnotacja.* Czas reakcji podawany jest w milisekundach;  $n$  – liczba obserwacji;  $M$  – średnia;  $SD$  – odchylenie standardowe;

**Tabela 23**

*Średnie i odchylenia standardowe dla średnich czasów reakcji (ms) w warunku niezgodności*

	Dzieci ( $n = 47$ )	Młodzież ( $n = 48$ )	Dorośli ( $n = 50$ )
Długość wyrazu	$M (SD)$	$M (SD)$	$M (SD)$
2 sylaby	1521 (216)	1132 (173)	963 (148)
3 sylaby	1607 (256)	1212 (197)	989 (135)

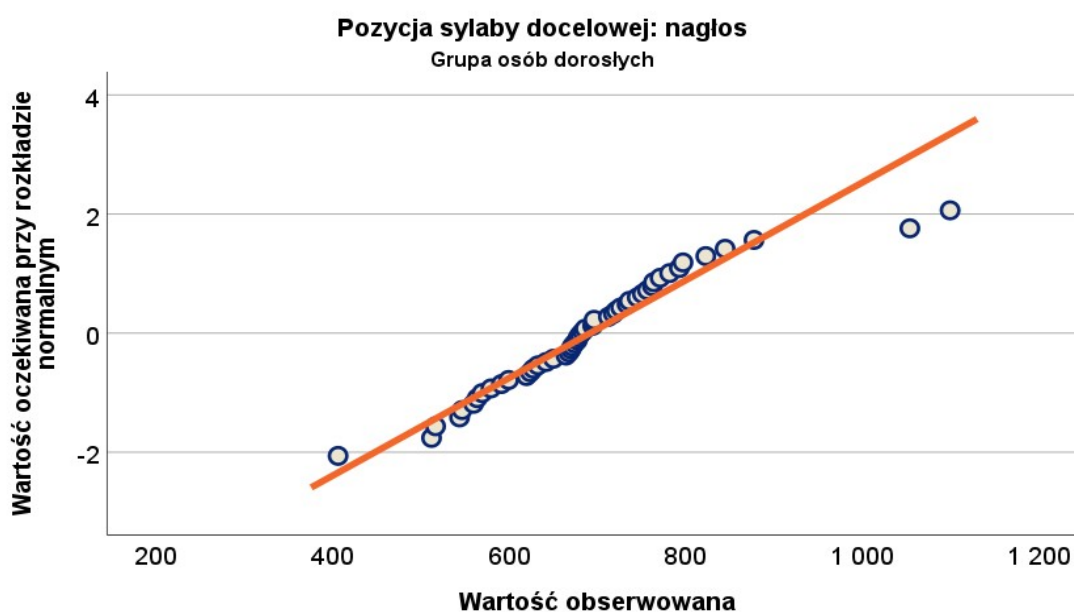
*Adnotacja.* Czas reakcji podawany jest w milisekundach;  $n$  – liczba obserwacji;  $M$  – średnia;  $SD$  – odchylenie standardowe;

<sup>162</sup> Do analizy włączono jedynie odpowiedzi osób, które dla każdej pozycji sylaby docelowej oraz obu długości wyrazów uzyskały minimum 30% poprawnych reakcji (zob. paragraf 5.4.2.1.2.)

Dokładniejszy ogląd rozkładu wyników w oparciu o wykresy K-K z trendem wykazał, że wyniki dwojga osób badanych z grupy dorosłych na tyle odstają od średniej, że zaburzają parametry rozkładu całej próby (zob. rysunek 21 oraz rysunek 22). Z tego względu postanowiono wykluczyć wyniki tych osób z porównań czasów reakcji dla warunku zgodności. Po ich usunięciu z bazy ponownie przeprowadzono test normalności Shapiro-Wilka, otrzymując tym razem wyniki nieistotne statystycznie w każdym z analizowanych pomiarów. Oznacza to brak wystarczających podstaw do odrzucenia hipotezy zerowej o zgodności analizowanych rozkładów z rozkładem normalnym.

### Rysunek 21

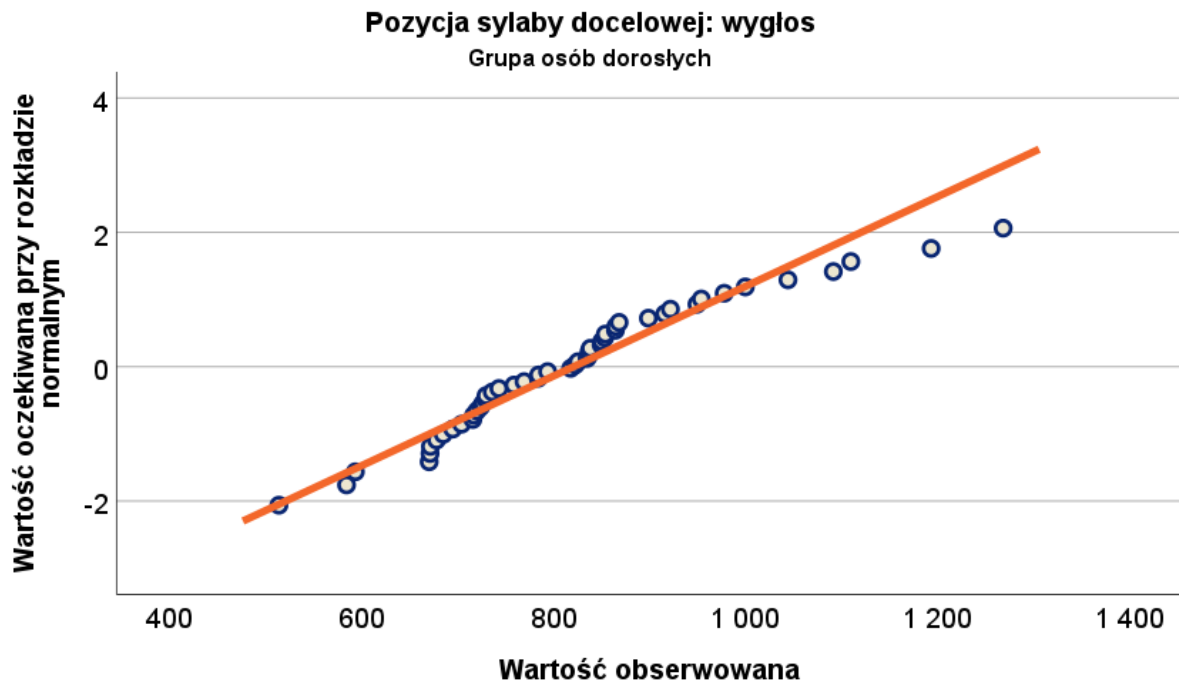
Wykres K-K z trendem dla rozkładu średnich czasu reakcji osób dorosłych na sylabę docelową w nagłosie



W związku z powyższym postanowiono hipotezy 1.2.2-1.2.4. zweryfikować przy pomocy dwóch dwuczynnikowych analiz wariancji w schemacie mieszanym. Pierwsza analiza wariancji poświęcona była wpływowi czynnika wieku (przynależność do grupy wiekowej) oraz pozycji wykrywanej sylaby na czas decyzji pozytywnej (wykrycie sylaby) w warunku zgodności w zadaniu *Monitorowanie sylab*. Druga analiza wariancji poświęcona była również wpływowi wieku (przynależność do grupy wiekowej) oraz długości wyrazu na czas poprawnej decyzji negatywnej (niewykrycie sylaby) w warunku niezgodności w zadaniu *Monitorowanie sylab*.

## Rysunek 22

Wykres K-K z trendem dla rozkładu średnich czasu reakcji osób dorosłych na sylabę docelową w wygłosie



### 6.3.1 Test hipotezy 1.2.1. Monitorowanie wewnętrzne zwiększa swoją dokładność wraz z wiekiem.

Mając na względzie zastrzeżenia dotyczące rozkładu poprawności odpowiedzi w grupach wiekowych, aby zweryfikować czy monitorowanie wewnętrzne zwiększa swoją efektywność wraz z wiekiem, postanowiono przeprowadzić jednoczynnikową analizę wariancji w schemacie międzygrupowym porównując ze sobą poprawność wyników uzyskanych przez uczestników badania z trzech grup wiekowych. Otrzymano wynik:  $F(2, 148) = 53,46$ ;  $p < 0,001$ ;  $\eta_p^2 = 0,419$ <sup>163</sup>. Porównania *post hoc* przeprowadzone za pomocą testu Tamhane'a ujawniły istotne różnice między każdą z grup badanych. Przedstawia je tabela 24. Na tej podstawie można stwierdzić, że wraz z wiekiem monitorowanie wewnętrzne zwiększa swoją efektywność. Tym samym można uznać, że **hipoteza 1.2.1. została potwierdzona**.

<sup>163</sup> Ze względu na niespełnienie założenia o homogeniczności wariancji dla międzyobiektowej zmiennej niezależnej (wartość testu Levene'a równości wariancji błędów była istotna statystycznie) zastosowano poprawkę Welcha. Wartości testu  $F$  oraz  $\eta_p^2$  podane są już w postaci skorygowanej.

**Tabela 24**

*Wyniki porównań parami (post hoc) między grupami wiekowymi pod względem poprawności odpowiedzi w zadaniu Monitorowanie sylab*

Zestawiane grupy wiekowe (A-B)	Różnica średnich	SE	p
Dorośli - dzieci	0,58	0,058	<0,001
Dorośli - młodzież	0,27	0,044	<0,001
Młodzież - dzieci	0,31	0,058	<0,001

*Adnotacja:* Zestawiano średni odsetek poprawnych odpowiedzi w poszczególnych grupach wiekowych;

SE – błąd standardowy; p – wartość p dla testu Tamhane’a z poprawką Bonferroniego dla wielokrotnych porównań

### **6.3.2 Test hipotez 1.2.2.-1.2.4. dotyczących rozwojowych oraz językowych uwarunkowań szybkości monitorowania wewnętrznego**

Testowane w tej sekcji hipotezy zostały sformułowane następująco:

- Hipoteza 1.2.2. Szybkość monitorowania wewnętrznego wzrasta wraz z wiekiem.
- Hipoteza 1.2.3. Szybkość monitorowania wewnętrznego maleje wraz z zajmowaniem przez wykrywany element kolejnych pozycji w strukturze wyrazu (wydłuża się czas efektywnego monitorowania).
- Hipoteza 1.2.4. Szybkość monitorowania wewnętrznego maleje wraz ze wzrostem długości monitorowanego wyrazu (wydłuża się czas efektywnego monitorowania).

Aby odpowiedzieć na pytanie, czy czas efektywnego monitorowania wewnętrznego jest uzależniony od wieku osób badanych oraz od pozycji sylaby docelowej w monitorowanym wyrazie, przeprowadzono dwuczynnikową analizę wariancji w schemacie mieszanym: 3 (grupa wiekowa) x 3 (pozycja sylaby). Czynnikiem mierzonym wewnątrz osób była pozycja sylaby w strukturze wyrazu, natomiast czynnikiem między osobami była przynależność do jednej z grup wiekowych. Zmienną zależną był czas efektywnego monitorowania wewnętrznego wyznaczony przez czas potrzebny na podjęcie decyzji o występowaniu sylaby docelowej w strukturze nazwy wyświetlanego obrazka (czas od wyświetlenia obrazka do uderzenia w klawisz klawiatury). Analiza wariancji została

przeprowadzona w modelu jednozmiennowym (*univariate repeated measures ANOVA*).

Uzyskano następujące rezultaty:

1. Istotny efekt główny wieku (przynależności do grupy wiekowej)  $F(2, 131) = 118,53$ ;  $p < 0,001$ ;  $\eta_p^2 = 0,644$ . Pozwala to uznać, że od przynależności do grupy wiekowej zależy ponad 60% zmienności w szybkości reagowania na rozpoznanie sylaby w monitorowanym wyrazie. Porównania *post hoc* za pomocą testu Scheffègo wykazały istotne różnice między wszystkimi grupami. Oznacza to, że w grupie dorosłych czas potrzebny na efektywne monitorowanie wewnętrzne był najniższy, a w grupie dzieci najwyższy. Młodzież istotnie szybciej monitorowała nazwy obrazków w porównaniu z osobami w wieku późnego dzieciństwa, jednocześnie przeznaczając na ten proces istotnie więcej czasu niż osoby dorosłe.
2. Istotny efekt główny pozycji sylaby  $F(2, 262) = 187,84$ ;  $p < 0,001$ ;  $\eta_p^2 = 0,589$ . Oznacza to, że prawie 60% zmienności w czasie monitorowania wynika z pozycji, jaką w monitorowanym wyrazie zajmuje sylaba docelowa. Analiza *post hoc* z korektą dla porównań wielokrotnych Bonferroniego wykazała istotne różnice ( $p < 0,001$ ) między czasem wykrywania sylaby w nagłosie, a wykrywaniem sylab na pozostałych pozycjach wyrazu. Jednocześnie czas monitorowania sylab w śródgłosie i wygłosie nie różnił się w sposób istotny ( $p = 0,15$ ). Wynik ten pozwala to na wyciągnięcie wniosku, że efektywny czas monitorowania jest uzależniony od pozycji sylaby docelowej i znacznie wzrasta między nagłosem i śródgłosem, utrzymując podobny poziom dla pozycji śródgłosu i wygłosu.
3. Istotny efekt interakcji przynależności do grupy wiekowej i pozycji sylaby  $F(4, 262) = 10,325$ ;  $p < 0,001$ ;  $\eta_p^2 = 0,136$ . Oznacza to, że efekt pozycji sylaby docelowej jest inny w poszczególnych grupach wiekowych. Analiza prostego efektu głównego pozycji sylaby pokazuje jednak, że nie mamy tu do czynienia z odmiennymi trendami w poszczególnych grupach, ale raczej z różną siłą efektu pozycji sylabowej, co obrazuje rysunek 23. Porównana parami (z korektą istotności Bonferroniego) pokazały, że w każdej z grup wiekowych różnica w czasie efektywnego monitorowania była istotna między nagłosem, a śródgłosem oraz nagłosem i wygłosie, ale nieistotna dla pozycji sylaby docelowej w śródgłosie i wygłosie. Pokazuje to, że w przypadku grupy dzieci, efekt różnicy czasu reakcji między nagłosem i pozostałymi pozycjami był największy<sup>164</sup>.  
Dokładne wyniki porównań przedstawia tabela 25.

---

164 Analiza prostego efektu głównego przynależności do grupy wiekowej wykazała, że największe różnice w średnich czasach reakcji między dziećmi i pozostałymi grupami wiekowymi dotyczyły pozycji śródgłosu, natomiast największa różnica w czasach reakcji między młodzieżą a dorosłymi wystąpiła w próbach z pozycją sylaby docelowej w wygłosie.

## Rysunek 23

Czas reakcji w zadaniu Monitorowanie sylab w zależności od pozycji sylaby docelowej w wyrazie: porównanie grup wiekowych

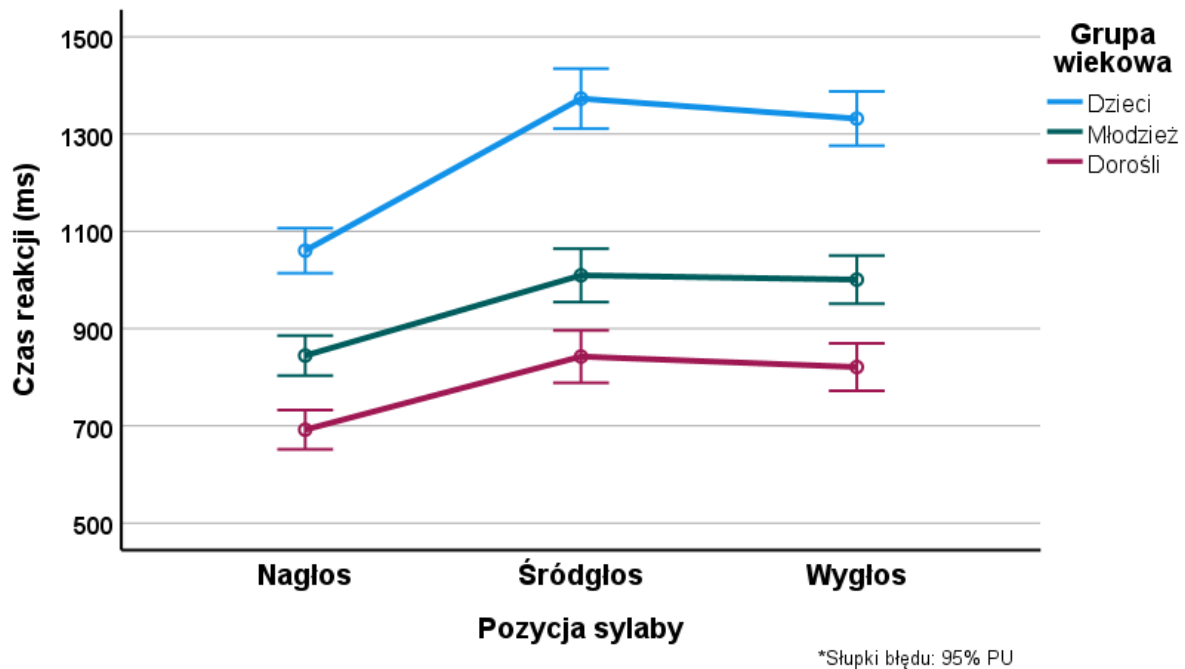


Tabela 25

Analiza różnic w średnim czasie reakcji na sylabę docelową w zależności od pozycji w wyrazie i grupy wiekowej

Grupa wiekowa	Zestawiane pozycje sylaby docelowej	Różnica średnich	SE	p
Dzieci	Śródgłos-nagłos	313**	23	<0,001
	Śródgłos-wygłos	41	23	0,240
	Wygłos-nagłos	271**	20	<0,001
Młodzież	Śródgłos-nagłos	165**	21	<0,001
	Śródgłos-wygłos	9	21	1,000
	Wygłos-nagłos	156**	17	<0,001
Dorośli	Śródgłos-nagłos	152**	21	<0,001
	Śródgłos-wygłos	24	21	0,742
	Wygłos-nagłos	128**	17	<0,001

Adnotacja: Zestawiano średnie czasy reakcji (wartość w milisekundach) w poszczególnych grupach wiekowych; SE – błąd standardowy; p – wartość p dla testów *post hoc* z poprawką Bonferroniego dla wielokrotnych porównań; \* $p < 0,05$ ; \*\* $p < 0,01$

Analogicznie do opisanej powyżej analizy, aby odpowiedzieć na pytanie, czy czas efektywnego monitorowania wewnętrznego jest uzależniony od wieku osób badanych (co zostało już wykazane w poprzedniej analizie) oraz od długości monitorowanego wyrazu (w przypadku negatywnego efektu monitorowania, tj. niewykrycia w strukturze wyrazu sylaby docelowej), przeprowadzono dwuczynnikową analizę wariancji w schemacie mieszanym: 3 (grupa wiekowa) x 2 (długość wyrazu). Czynnikiem mierzonym wewnątrz osób była długość wyrazu (2 vs. 3 sylaby), natomiast czynnikiem między osobami była przynależność do jednej z grup wiekowych. Zmienną zależną był czas efektywnego monitorowania wewnętrznego wyznaczony przez czas potrzebny na podjęcie decyzji o niewystępowaniu sylaby docelowej w strukturze nazwy wyświetlanego obrazka (czas od wyświetlenia obrazka do uderzenia w klawisz klawiatury). Analiza wariancji została przeprowadzona w modelu jednozmiennowym (*univariate repeated measures ANOVA*). Otrzymano następujące wyniki:

1. Istotny efekt główny przynależności do grupy wiekowej  $F(2, 142) = 122,11$ ;  $p < 0,001$ ;  $\eta_p^2 = 0,632$ <sup>165</sup>. Wynik ten pozwala ponownie uznać, że czas w jakim osoby badane monitorowały nazwy obrazków zależał w dużej mierze (ponad 60% wyjaśnionej wariancji) od ich wieku. Porównania *post hoc* za pomocą testu Tamhane'a wykazały, że istotne różnice występowały między wszystkimi grupami. Oznacza to, że w grupie dzieci czas potrzebny na efektywne monitorowanie wewnętrzne był najwyższy, a w grupie dorosłych – najniższy. Młodzież istotnie szybciej monitorowała nazwy obrazków w porównaniu z osobami w wieku późnego dzieciństwa, jednocześnie przeznaczając na ten proces istotnie więcej czasu niż osoby dorosłe.
2. Istotny efekt główny długości wyrazu  $F(1, 142) = 37,16$ ;  $p < 0,001$ ;  $\eta_p^2 = 0,207$ . Na podstawie tego wyniku można uznać, że szybkość monitorowania u osób w różnym wieku jest uzależniona częściowo od długości analizowanego materiału: wyrazy krótsze (2-sylabowe) monitorowane są szybciej niż wyrazy dłuższe (3-sylabowe).
3. Istotny efekt interakcji czynnika długości wyrazu i wieku (przynależność do grupy wiekowej)  $F(2, 142) = 3,34$ ;  $p = 0,038$ ;  $\eta_p^2 = 0,045$ . Oznacza to, że różnica w długości wyrazu w różnym stopniu wpływa na szybkość monitorowania w poszczególnych grupach wiekowych, choć efekt należy uznać za słaby (Ferguson, 2009; por. Niewiarowski i in., 2013). Analiza prostego efektu głównego długości wyrazu w poszczególnych grupach wiekowych wykazała, że w grupie dzieci różnica między monitorowaniem wyrazów 2- i 3-sylabowych była większa niż w pozostałych grupach

---

<sup>165</sup> Ze względu na niespełnienie założenia o homogeniczności wariancji dla międzyobiektowej zmiennej niezależnej (wartość testu Levene'a równości wariancji błędów była na granicy przyjętej istotności statystycznej) zastosowano poprawkę Welcha. Wartości testu  $F$  oraz  $\eta_p^2$  podane są już w postaci skorygowanej.

wiekowych, natomiast u osób dorosłych była ona na tyle niewielka, że okazała się nieistotna statystycznie ( $p > 0,05$ ). Wyniki porównań parami zawiera tabela 26.

Graficzne podsumowanie wyników przeprowadzonej analizy wariancji przedstawia wykres na rysunku 24.

**Tabela 26**

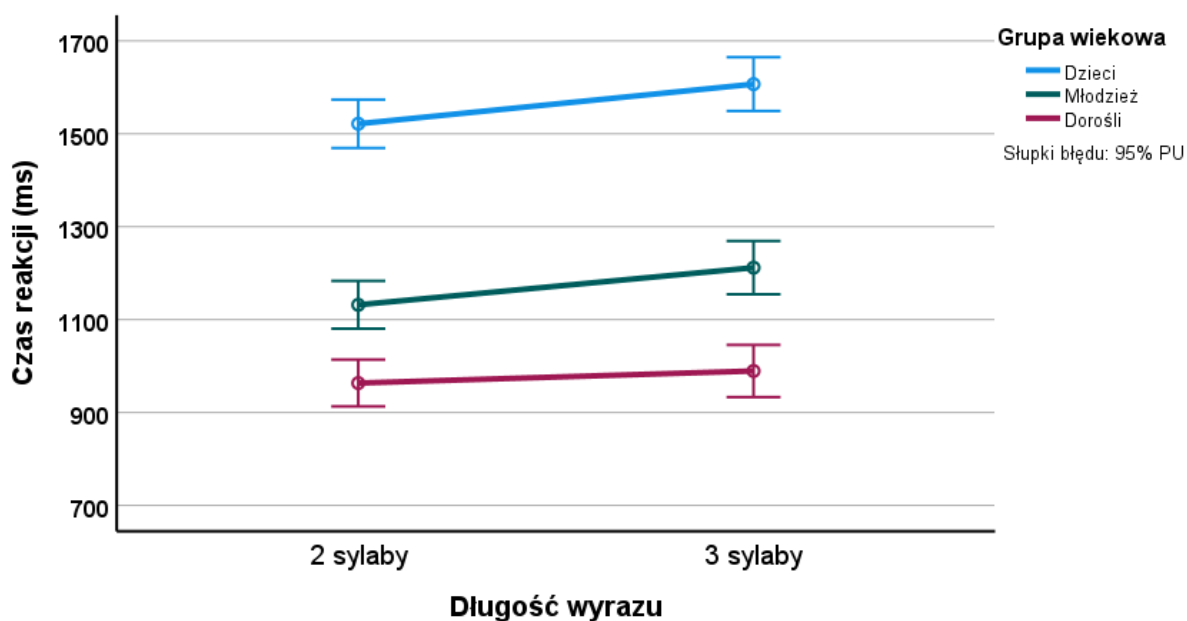
*Analiza różnic w czasie monitorowania sylab między wyrazami 2- i 3-sylabowymi w zależności od wieku uczestników*

Grupa wiekowa	Różnica średnich czasu monitorowania wyrazów 3- i 2-sylabowych	SE	p
Dzieci	85,41**	18,38	0,001
Młodzież	79,98**	18,18	0,001
Dorośli	26,01	17,82	0,147

*Adnotacja:* Zestawiano średnie czasy reakcji (wartość w milisekundach) w poszczególnych grupach wiekowych; SE – błąd standardowy; p – wartość p dla testów post hoc z poprawką Bonferroni dla wielokrotnych porównań; \* $p < 0,05$ ; \*\* $p < 0,01$

**Rysunek 24**

*Czas reakcji w zadaniu Monitorowanie Sylab w zależności od długości wyrazu: porównanie grup wiekowych*



Podsumowując wyniki obu analiz wariancji przeprowadzonych na wynikach czasów reakcji (poprawnych odpowiedzi) w zadaniu *Monitorowanie sylab*, można uznać, że:

1. Wiek jest istotnym czynnikiem warunkującym szybkość monitorowania sylab w generowanych wyrazach. Zarówno w warunku zgodności, kiedy należało udzielić odpowiedzi TAK, jak i w warunku niezgodności, kiedy poprawną decyzją był wybór odpowiedzi NIE dorośli najszybciej osiągnęli poprawne rezultaty monitorowania, w porównaniu do młodzieży i dzieci. Z kolei osoby w wieku późnego dzieciństwa potrzebowały na przebieg procesu monitorowania najwięcej czasu w porównaniu z przedstawicielami starszych grup wiekowych. W związku z powyższym, **hipotezę 1.2.2. mówiącą iż, czas efektywnego monitorowania wewnętrznego skraca się wraz z wiekiem, można uznać za potwierdzoną.**
2. Pozycja sylaby docelowej w strukturze monitorowanego wyrazu ma znaczenie dla czasu efektywnego monitorowania. Na podstawie przeprowadzonej analizy statystycznej wykazano, że w odpowiedziach osób z każdej grupy wiekowej wykrywanie sylab na pierwszej pozycji (nagłos) zajmowało istotnie mniej czasu niż wykrywanie sylab na dalszych pozycjach (śródgłos i wygłos). Wbrew przewidywaniom, różnica czasów reakcji między śródgłosem i wygłosem nie była istotna statystycznie. Z tego względu **hipoteza 1.2.3. mówiąca o tym, że czas efektywnego monitorowania wewnętrznego wydłuża się wraz z zajmowaniem przez wykrywany element kolejnych pozycji w strukturze wyrazu, uzyskała jedynie częściowe potwierdzenie:** efekt ten widoczny jest jedynie w porównaniu pozycji nagłosowej z pozostałymi pozycjami w strukturze wyrazu 3-sylabowego.
3. Długość monitorowanego wyrazu ma znaczenie dla czasu efektywnego monitorowania sylab. Na podstawie przeprowadzonej analizy statystycznej wykazano, że monitorowanie wyrazów dwusylabowych zajmowało istotnie mniej czasu niż monitorowanie wyrazów trójsylabowych. Wpływ długości wyrazu był jednak uzależniony od wieku osób badanych, gdyż najsilniej ujawnił się w grupie dzieci, natomiast nie był istotny statystycznie w grupie osób dorosłych. W związku z tym można uznać, że **hipoteza 1.2.4., mówiąca o tym, że czas efektywnego monitorowania wewnętrznego wydłuża się wraz ze wzrostem długości monitorowanego wyrazu uzyskała zasadnicze wsparcie w wynikach badania,** ale jednocześnie efekt ten może mieć marginalne znaczenie dla automonitorowania produkcji mowy przez osoby dorosłe, przy niewielkich różnicach w długości wyrazu (różnica jednej sylaby).

4. Siła z jaką strukturalne uwarunkowania tekstu (długość wyrazu oraz pozycja docelowej sylaby w wyrazie) wpływają na szybkość monitorowania jest częściowo uwarunkowana przez wiek osób badanych. Istotne efekty interakcji czynników głównych w obu analizach wariancji oraz dalsze analizy prostych efektów głównych wykazały, że w odpowiedziach osób dorosłych efekt pozycji sylaby oraz długości wyrazu były najmniejsze ze wszystkich grup wiekowych (efekt długości wyrazu był nieistotny statystycznie). Pozwala to na wyciągnięcie wniosku, że wraz z wiekiem maleje znaczenie czynników strukturalnych tekstu dla szybkości efektywnego monitorowania. Rezultat ten można uznać za wspierający hipotezę o ogólnym wzroście sprawności automonitorowania procesu mówienia wraz z wiekiem (hipoteza 1.2.).

## 6.4 Reperacje błędów

Kolejne dwie testowane hipotezy dotyczyły zmian rozwojowych w zakresie dokonywania reperacji. Sposób korygowania błędów symulowany był w zadaniu *Powtarzania zdań*, które – ze względu na manipulację miejscem błędu do poprawienia – dostarczyło dwojakiego rodzaju danych:

- a) szybkość dokonywania interupcji w błędnych wypowiedziach po wykryciu błędu przez monitorowanie zewnętrzne (warunek błędnego podmiotu);
- b) odsetek reperacji ukrytych względem innych reperacji (warunek błędnego dopełnienia)

Aby zweryfikować hipotezę na temat różnic w szybkości dokonywania interupcji u osób w odrębnych grupach wiekowych przeanalizowano średnie długości wypowiedzi z reperandum (średnia liczba wypowiedzianych sylab od momentu rozpoczęcia powtarzania zdania do momentu rozpoczęcia reperacji), wyznaczające moment przerwania wypowiedzi (moment interupcji) w warunku błędnego podmiotu. Tabela 27 przedstawia średnie, odchylenia standardowe i inne statystyki tej zmiennej z podziałem na poszczególne grupy wiekowe.

Test Shapiro-Wilka dla każdej próby zwrócił wynik istotny statystycznie ( $p < 0,05$ ), co nie pozwala na uznanie rozkładów wyników w grupach za zgodne z rozkładem normalnym. Tym samym porównania międzygrupowe na potrzeby weryfikacji hipotezy 1.3.1. możliwe są przy zastosowaniu testów nieparametrycznych. W warunku błędnego dopełnienia zbierane były natomiast dane o częstości reperacji ukrytych w strategii odroczenia na tle innych reperacji możliwych do zastosowania zgodnie z procedurą zadania. Dane te miały posłużyć do weryfikacji hipotezy na temat wzrastającego wraz z wiekiem odsetka reperacji ukrytych. Tabela 28 przedstawia statystyki opisowe dla tej zmiennej.

**Tabela 27**

*Statystyki opisowe dla długości wypowiedzi reparamandum (liczba sylab) w warunku błędnego podmiotu*

Grupa wiekowa	<i>M</i>	<i>SD</i>	<i>Me</i>	<i>SKE</i>	<i>K</i>	<i>W</i>	<i>p</i>
Dzieci ( <i>n</i> = 50)	4,63	1,54	4,38	0,275	-1,176	0,93	0,007
Młodzież ( <i>n</i> = 50)	4,86	1,2	4,69	0,508	-0,723	0,94	0,015
Dorośli ( <i>n</i> = 50)	4,77	1,02	4,51	0,566	-0,492	0,95	0,033

*Adnotacja.* *n* – liczba obserwacji; *M* – średnia; *SD* – odchylenie standardowe; *Me* – mediana, *SKE* – skośność; *K* – kurtoza; *W* – wynik testu Shapiro-Wilka; *p* – wartość *p* dla testu Shapiro-Wilka

**Tabela 28**

*Statystyki opisowe dla procenta odpowiedzi z reperacją ukrytą (przez odroczenie) w warunku błędnego dopełnienia w trzech grupach wiekowych*

Grupa wiekowa	<i>M</i>	<i>SD</i>	<i>Me</i>	<i>SKE</i>	<i>K</i>	<i>W</i>	<i>p</i>
Dzieci ( <i>n</i> = 49)	50	32	50	-0,249	-1,130	0,93	0,005
Młodzież ( <i>n</i> = 50)	58	29	62,5	-0,607	-0,439	0,93	0,007
Dorośli ( <i>n</i> = 50)	72	26	81	-0,916	-0,233	0,88	0,0001

*Adnotacja.* Tabela przedstawia wartości procentowe; *n* – liczba obserwacji; *M* – średnia; *SD* – odchylenie standardowe; *Me* – mediana, *SKE* – skośność; *K* – kurtoza; *W* – wynik testu Shapiro-Wilka; *p* – wartość *p* dla testu Shapiro-Wilka

Wyniki jednego dziecka wykluczono z analizy z powodu poprawności odpowiedzi w warunku błędnego dopełnienia poniżej 30%. Z danych w tabeli wynika, że dzieci średnio w 50% odpowiedzi poprawiały powtarzane zdania poprzez reperację ukrytą wypowiedzi (w tych wypadkach powtarzane błędne zdanie od razu otrzymywało poprawne, tj. zgodne z obrazkiem dopełnienie bez konieczności odtwarzania części reparamandum). W starszych grupach wiekowych odpowiedzi takich było więcej, średnio 58% w grupie młodzieży i 72% w grupie

dorosłych. Z racji tego, że dane przedstawiają proporcje (odsetek odpowiedzi, w których dokonano reperacji ukrytej), czyli mieszczą się w zakresie 0-1 i nie spełniają założenia o normalności rozkładów w poszczególnych grupach, do weryfikacji hipotezy 1.3.2. postanowiono zastosować nieparametryczny test statystyczny.

#### **6.4.1 Test hipotezy 1.3.1. Szybkość dokonywania interupcji w błędnej wypowiedzi wzrasta wraz z wiekiem.**

Dane w tabeli 27 pokazują, że średnia długość wypowiedzi z reparamum (liczba wypowiedzianych sylab przed wprowadzeniem reperacji), a tym samym moment interupcji są bardzo zbliżone we wszystkich grupach wiekowych. Aby sprawdzić, czy różnice te są jednak istotne statystycznie przeprowadzono test  $H$  Kruskala-Wallisa dla prób niezależnych. Uzyskano nieistotny statystycznie wynik porównania długości reparamum:  $H(2, N = 150) = 1,57$ ;  $p = 0,457$ ;  $\eta^2_H = -0,003$ , zatem nie ma podstaw do odrzucenia hipotezy zerowej o równości rozkładów. Oznacza to, że dzieci, młodzież i dorośli przerywali powtarzaną wypowiedź po wypowiedzeniu podobnej liczby sylab (średnio między 4. a 5. sylabą od rozpoczęcia wypowiedzi). **Na podstawie uzyskanego wyniku nie można uznać hipotezy 1.3.1. za potwierdzoną.**

Jednocześnie analiza wizualna wykresu na rysunku 25 pokazuje, że rozkład wyników w grupie dzieci charakteryzował się największą wariancją, a większa niż w innych grupach niejednorodność wyników pozwala sądzić, że istnieje inny czynnik, który w dużym stopniu może odpowiada za tę zmienność. Aby dokładniej zrozumieć zależności między uzyskanymi wynikami, a wiekiem, postanowiono wykonać dodatkowe analizy:

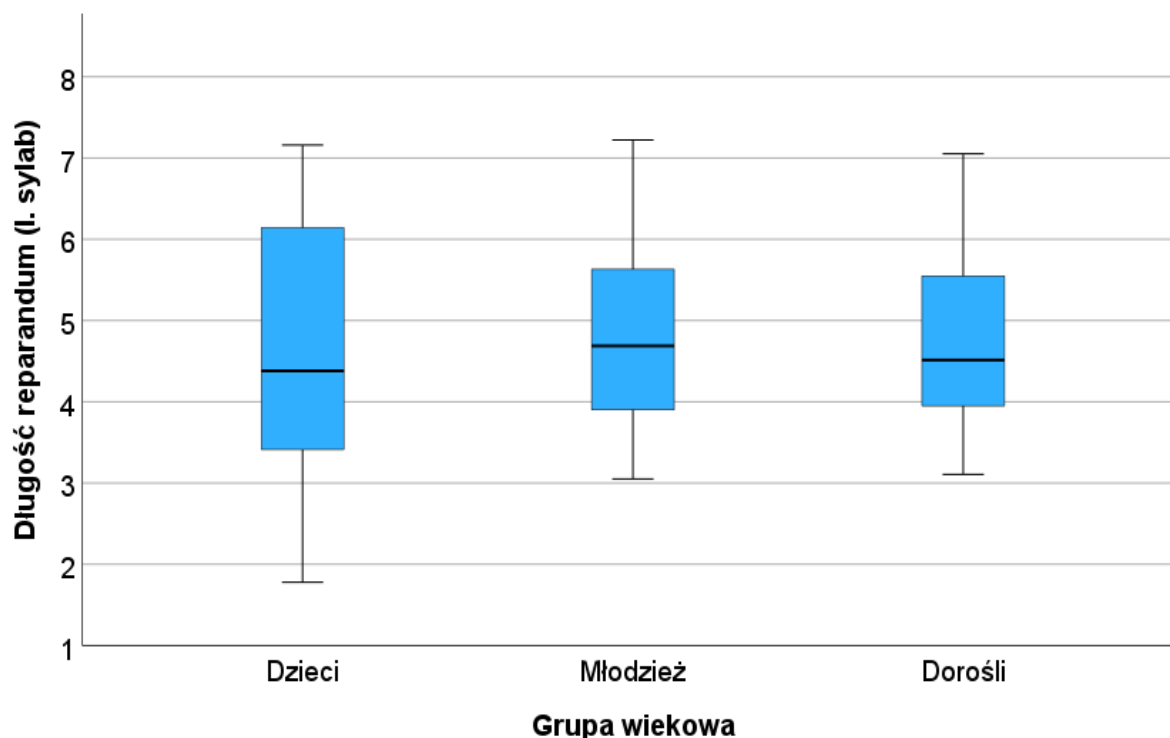
1. Porównanie procentowego rozkładu odpowiedzi z interupcją w każdej z grup wiekowych.
2. Sprawdzenie (a) różnic w tempie mówienia (tempie wypowiedzi)<sup>166</sup> między poszczególnymi grupami oraz (b) związku tempa mówienia z momentem dokonywania interupcji.
3. Porównanie międzygrupowo bezwzględnego czasu reakcji (przerwania wypowiedzi w celu wprowadzenia reperacji).
4. Korelacja między momentem zatrzymania mówienia, a wiekiem przy kontroli tempa mówienia.

---

<sup>166</sup> Jak wspomniano w sekcji 5.4.2.1.3. zastosowana procedura obliczania tempa mówienia (por. Frydrychowicz, 1999) obejmowała cały czas trwania wypowiedzi (łącznie z pauzami), dlatego uzyskana miara odpowiada temu, co w literaturze (Cholewiak, 2016; Michalik i in., 2018) określane jest jako tempo wypowiedzi.

## Rysunek 25

Porównanie rozkładów średniej długości wypowiedzi z reparamdum w trzech grupach wiekowych



### 6.4.1.1 Dodatkowa analiza (1): Porównanie procentowego rozkładu odpowiedzi z interupcją w każdej z grup wiekowych.

Na potrzeby weryfikacji hipotezy 1.3.1 przyjęto, że szybkość dokonywania interupcji liczona będzie w postaci liczby sylab wypowiedzianych od rozpoczęcia zdania do momentu wprowadzenia korekty<sup>167</sup> (długość wypowiedzi z reparamdum). Ma ona tę zaletę, że pozwalała na objęcie jednym wskaźnikiem odpowiedzi, które zostały przerwane przed reperacją wraz z tymi, w których reparamdum składało się z całego powtarzanego zdania (zdanie powtórzone w całości bez interupcji), a korekta dodawana była już po wypowiedzeniu ostatniego wyrazu. Miara ta nie różnicuje natomiast faktycznej liczby wypowiedzi przerwanych względem pełnych powtórzeń prezentowanego w słuchawkach zdania. Nie pozwala zatem na wykazanie, czy w którejś z badanych grup tendencja do przerywania wypowiedzi była wyższa.

W związku z tym, że nie udało się wykazać różnic międzygrupowych, co do momentu przerywania wypowiedzi, postanowiono sprawdzić, czy osoby w badanych grupach wiekowych różnią się pod względem liczby przerwanych wypowiedzi. Inaczej mówiąc,

<sup>167</sup> Podobnym wskaźnikiem (opartym na liczbie sylab) posłużył się Levelt (1983; 1989) wskazując dystans dzielący wyraz zawierający błąd od momentu interupcji.

postanowiono sprawdzić, czy – w nawiązaniu do niepotwierdzonej hipotezy 1.3.1. – osoby dorosłe będą częściej przerywać błędne wypowiedzi w porównaniu do młodzieży i dzieci. Przyjęto zatem robocze założenie, że zdolność do zahamowania procesu mówienia wymaga kontroli, która rozwija się wraz z wiekiem, co może ujawnić się także w częstości dokonywania interupcji (nawet jeśli nie wykazano istotnych różnic międzygrupowych szybkości dokonywania interupcji).

Zestawienie w tabeli 29 odsetka przerwanych odpowiedzi średnio dla każdej grupy ujawnia pewne różnice. Dzieci przerywały średnio 66% wypowiedzi, młodzież 74%, a dorośli 80% wypowiedzi. Z racji tego, że dane przedstawiają proporcje (odsetek odpowiedzi, w których przerwano powtarzane zdanie przed jego ukończeniem), czyli mieszczą się w zakresie 0-1 i nie spełniają założenia o normalności rozkładów w poszczególnych grupach, nie jest więc możliwe przeprowadzenie porównań w oparciu o średnie i wykorzystanie parametrycznych testów statystycznych.

**Tabela 29**

*Procent odpowiedzi z interupcją w warunku błędnego podmiotu*

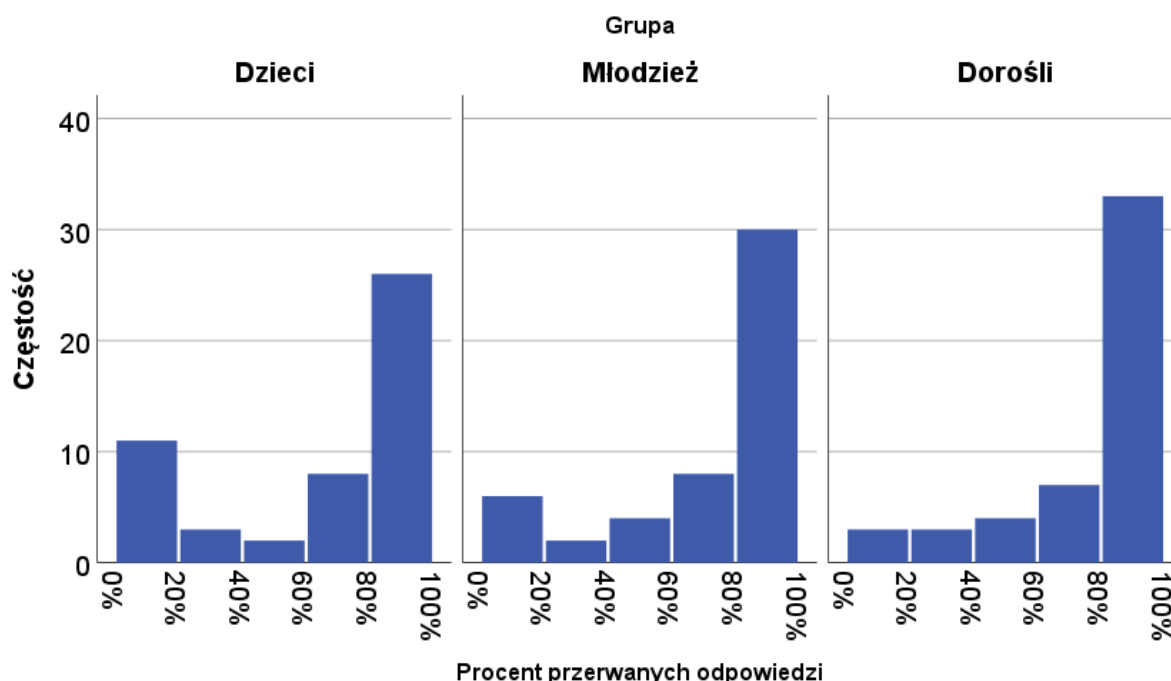
Grupa wiekowa	<i>M (SD)</i>	Me
Dzieci ( <i>n</i> = 50)	66 (37)	83
Młodzież ( <i>n</i> = 50)	74 (33)	91
Dorośli ( <i>n</i> = 50)	80 (27)	94

*Adnotacja.* Tabela przedstawia statystyki dla wartości procentowych; *n* – liczba obserwacji; *M* – średnia; *SD* – odchylenie standardowe; *Me* – mediana

Na rysunku 26 umieszczone zostały histogramy reprezentujące częstości występowania różnych proporcji odpowiedzi z interupcją w trzech grupach wiekowych. Choć rozkłady wyników są podobne we wszystkich grupach wiekowych, można zauważyć, że pierwszy przedział klasowy obejmujący osoby z najniższym odsetkiem przerwanych odpowiedzi jest najliczniejszy w grupie dzieci (11 osób), a najmniej liczny w grupie dorosłych (3 osoby). W grupie młodzieży w pierwszym przedziale klasowym znalazło się natomiast 6 osób. Odpowiednio ostatni przedział klasowy, reprezentujący liczebność osób z najwyższym odsetkiem przerwanych odpowiedzi jest najliczniejszy w grupie osób dorosłych (33 osoby), nieco mniej liczny w grupie młodzieży (30 osób) oraz najmniej liczny w grupie dzieci (26 osób).

## Rysunek 26

Rozkłady częstości występowania interupcji w odpowiedziach w warunku błędnego podmiotu: porównanie grup wiekowych



Pokazuje to możliwy trend rozwojowy, polegający na tendencji do częstszego przerywania odpowiedzi u osób starszych wiekiem, względem młodszych użytkowników mowy. Aby porównać rozkłady wyników (procent przerwanych odpowiedzi) w trzech grupach wiekowych i sprawdzić, czy sugerowany trend przejdzie weryfikację statystyczną, wykorzystano nieparametryczny test Jonckheere-Terpsty. Uzyskano wynik:  $T_{JT} = 4237$ ;  $z = 1,71$ ;  $p = 0,045^{168}$ ;  $\tau_b = 0,113$ .

Zgodnie z oczekiwaniami ujawnione różnice okazały się więc istotne statystycznie, choć trend wzrostowy jest bardzo słaby, a testy *post hoc* (z korektą Bonferroniego na wielokrotne porównania) nie wykazały istotnych różnic dla porównań w parach. Można więc uznać, że w każdej z grup wiekowych dominowały odpowiedzi z przerywaniem reparandum przed dokończeniem jego realizacji, a tendencja ta nieznacznie wzrastała w grupach reprezentujących kolejne okresy rozwojowe.

<sup>168</sup> Dwustronna (asymptotyczna) wartość  $p$  wyniosła 0,087, a jednostronna  $p$ -wartość uzyskana przy pomocy symulacji Monte Carlo wyniosła 0,045, spełniając tym samym kryterium zadanej istotności  $\alpha=0,05$ . Ze względu na założony jednokierunkowy trend rozwojowy, uzyskana w symulacji jednostronna wartość  $p$  poniżej 0,05 jest wystarczająca do odrzucenia hipotezy zerowej.

#### 6.4.1.2 Dodatkowa analiza (2): Różnice w tempie mówienia między poszczególnymi grupami, a długość wypowiedzi z reparamandum

W związku z brakiem wykazania różnic między grupami wiekowymi pod względem szybkości dokonywania interupcji (brak różnic w średniej długości wypowiedzi z reparamandum) postanowiono zweryfikować związki szybkości dokonywania interupcji oraz wieku osób badanych z inną zmienną, tj. tempem mówienia. Według cytowanej wcześniej literatury tempo mówienia wzrasta w dzieciństwie, osiągając najwyższy poziom w okresie adolescencji, po czym ulega obniżeniu (Sabin i in., 1979). Różnice w tempie mówienia mogłyby mieć wpływ na wynik zadania *Powtarzanie zdań* w tym sensie, że zgodnie z procedurą (zob. 5.4.2.1.3.) obrazek służący za bodziec do przerwania mówienia i przygotowania reperacji pojawiał się 150 milisekund po rozpoczęciu mówienia (150 ms od zarejestrowania przez klawisz głosowy dźwięku). Przy potencjalnych różnicach w tempie mówienia, ten krótki odcinek czasu od rozpoczęcia mówienia do wyświetlenia obrazka, mógłby nieznacznie faworyzować osoby mówiące wolniej. Aby zweryfikować przypuszczenie o roli tempa mówienia w wykonaniu zadania *Powtarzanie zdań* w warunku błędnego podmiotu, przeprowadzono trzy kolejne analizy:

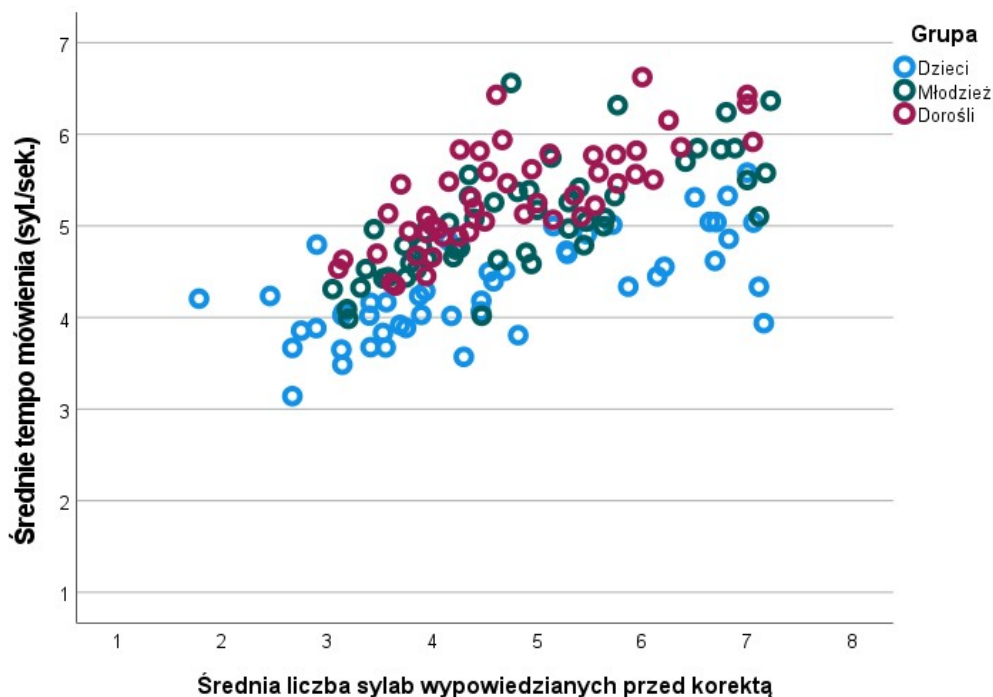
- a) wykonano porównanie tempa mówienia w warunku neutralnym (zgodność powtarzanego zdania z treścią obrazka) we wszystkich trzech grupach wiekowych poprzez przeprowadzenie jednoczynnikowej analizy wariancji w schemacie międzygrupowym (statystyki opisowe dla zmiennej „tempo mówienia” zawiera Załącznik F). Otrzymany wynik wykazał istotny wpływ czynnika wieku na tempo mówienia,  $F(2, 147) = 43,18$ ;  $p < 0,001$ ;  $\eta_p^2 = 0,37$ , co oznacza, że średnie wartości zmiennej zależnej różnią się istotnie między grupami wiekowymi. Stosunkowo duża wielkość efektu oznacza, że czynnik wieku tłumaczy 37% wariancji tempa mówienia. Porównania *post hoc* przy użyciu testu Scheffègo wykazały istotne różnice między grupą dzieci, a pozostałymi grupami wiekowymi. Różnica w tempie mówienia między młodzieżą a dorosłymi okazała się nieistotna statystycznie (zob. Załącznik F).
- b) wykonano analizę korelacji między tempem mówienia, a szybkością przerywania wypowiedzi (długością wypowiedzi z reparamandum). Wynik korelacji Pearsona wykazał istotną i silną zależność liniową między obiema zmiennymi:  $r = 0,61$ ;  $p < 0,001$ . Dodatni współczynnik korelacji oznacza, że wraz ze wzrostem tempa mówienia zwiększa się też liczba sylab wypowiedzianych od rozpoczęcia mówienia, do wprowadzenia korekty w zadaniu *Powtarzanie zdań*. Może to oznaczać, że osoby które mówią szybciej, z opóźnieniem wyhamowują swoją wypowiedź w porównaniu

do osób mówiących wolniej ale też, że od momentu rozpoczęcia mówienia do wyświetlenia obrazka i identyfikacji błędu w wypowiedzianym zdaniu zdążą wyartykułować większą jego część (więcej sylab) niż osoby mówiące wolniej.

- c) przeprowadzono analizę korelacji zmiennej „wiek” oraz średniej liczby sylab wypowiedzianych przed wprowadzeniem reperacji (długość wypowiedzi z reparandum) przy kontroli zmiennej „tempo mówienia”, która różnicuje osoby ze względu na wiek (analiza w podpunkcie a.) oraz wykazuje dodatnie związki ze średnią długością reparandum (analiza w podpunkcie b.). Analiza korelacji cząstkowej wykazała istnienie umiarkowanego związku między wiekiem i szybkością dokonywania interupcji (długością reparandum) przy kontroli wpływu tempa mówienia, jako kowariantu ( $r=-0,44$ ;  $p<0,001$ ). Jest to związek ujemny, co oznacza, że przy stałej wartości tempa mówienia wraz z wiekiem zmniejsza się liczba sylab wypowiedzianych przed wprowadzeniem reperacji (wraz z wiekiem wzrasta szybkość dokonywania interupcji w błędnej wypowiedzi, por. hipoteza 1.3.1.). Na rysunku 27 zamieszczono wykres rozrzutu ilustrujący opisywane zależności.

## Rysunek 27

Wykres rozrzutu dla relacji między średnim tempem mówienia, a szybkością dokonywania interupcji.



Podsumowując, powyższe wyniki świadczą o tym, że wraz z wiekiem zwiększa się tempo mówienia, które okazało się istotną zmienną (nieuwzględnioną wcześniej w modelu) związaną z szybkością dokonywania interupcji w zadaniu *Powtarzanie zdań*. Przy kontroli tej zmiennej okazało się, że wyrażone w hipotezie 1.3.1. przypuszczenie o wzrastającej wraz z wiekiem szybkości dokonywania interupcji (coraz mniej sylab wypowiedzi z reparamum przed wprowadzeniem reperacji) znajduje swoje potwierdzenie w analizie korelacji.

#### **6.4.1.3 Dodatkowa analiza (3): Porównanie bezwzględnego czasu reakcji (przerwania wypowiedzi w celu wprowadzenia reperacji) między grupami wiekowymi.**

Zastosowana pierwotnie miara szybkości dokonywania interupcji w postaci długości wypowiedzi z reparamum polegała na wyznaczeniu miejsca w powtarzonym zdaniu, w którym było ono przerywane, aby wprowadzić reperację (moment interupcji). Nie były natomiast zbierane dane na temat bezwzględnego czasu między rozpoczęciem mówienia, a jego przerwaniem i wprowadzeniem korekty (lub wypowiedzeniem w całości błędnego zdania i wprowadzeniem reperacji). Jeśli – jak wynika z powyższych analiz – młodzież i dorośli mówią w szybszym tempie niż dzieci, ale przerywają swoje wypowiedzi w podobnym momencie (średnio po wypowiedzeniu 4-5 sylab), to możliwe, że bezwzględny czas od rozpoczęcia mówienia do interupcji był w grupie dzieci dłuższy niż w pozostałych grupach. Aby sprawdzić tę ewentualność, obliczono<sup>169</sup> przybliżony średni czas trwania reparamum dla każdej z badanych osób na podstawie tempa mówienia (w warunkach neutralnych) oraz liczby sylab wypowiedzianych do momentu wprowadzenia korekty (w warunkach błędnego podmiotu).

Przeprowadzona jednoczynnikowa analiza wariancji w schemacie międzygrupowym wykazała istotne różnice między grupami,  $F(2, 147) = 7,98; p < 0,001; \eta_p^2 = 0,098$ . Niska siła efektu wskazuje, że wiek w niewielkim stopniu wyjaśnia różnice w szybkości przerywania mówienia po wykryciu błędu. Porównania *post hoc* z wykorzystaniem testu Tamhane'a wykazały, że jedynie różnica między dziećmi i dorosłymi jest istotna statystycznie na poziomie  $p < 0,001$ . Różnice między grupami dzieci i młodzieży oraz młodzieży i dorosłych nie uzyskały poziomu zadanej istotności statystycznej (odpowiednio:  $p = 0,072$  oraz  $p = 0,111$ ). Można więc powiedzieć, że dorośli nieco szybciej niż dzieci reagowali na wykrycie błędu w zdaniu, zatrzymując swoją wypowiedź średnio o 0,17 sek. wcześniej niż

---

<sup>169</sup> Jak wspomniano, dokładne dane o czasie trwania wypowiedzi z reparamum nie były zbierane podczas opracowania danych. Rejestrowano liczbę sylab w wypowiedzi oraz całkowity czas jej trwania. Na podstawie tych dwóch wartości obliczano tempo mówienia (liczba sylab na sekundę) w każdej próbie. Por. sekcja 5.4.2.1.3.

dzieci, co przy istotnej różnicy w tempie mówienia, skutkowało przerwaniem wypowiedzi w podobnym miejscu zdania (po około 4-5 sylabach).

Podsumowując, **wyniki dostarczone przez zadanie *Powtarzania zdań*, nie dają wystarczających podstaw do bezpośredniego potwierdzenia hipotezy 1.3.1. w oparciu o przyjęty wskaźnik szybkości dokonywania interupcji** i uznania, że szybkość ta wzrasta wraz z wiekiem. Dodatkowa analiza odpowiedzi pozwala jednak stwierdzić, że badane osoby dorosłe częściej niż dzieci przerywały powtarzane zdania, nie doprowadzając błędnej wypowiedzi do końca (wykazany trend rozwojowy we wzroście odsetka przerywanych odpowiedzi wraz z wiekiem). Ponadto, dalsze analizy wykazały istotną rolę tempa mówienia w szybkości dokonywania interupcji. Osoby, które mówiły generalnie szybciej, przed dokonaniem korekty wypowiadały więcej sylab błędnej wypowiedzi niż osoby o wolniejszym tempie mówienia. Z racji tego, że tempo mowy dorosłych (a także młodzieży) było istotnie wyższe niż tempo mówienia dzieci, do momentu wprowadzenia korekty wypowiadali oni podobną liczbę sylab, mimo iż u dzieci czas od rozpoczęcia mówienia do jego przerwania (*error-to-cutoff*) trwał istotnie dłużej o średnio prawie 0,2 sek. niż u dorosłych. **Dodatkowe analizy nie pozwalają więc na jednoznaczne odrzucenie hipotezy 1.3.1.**, dostarczając argumentów na jej poparcie.

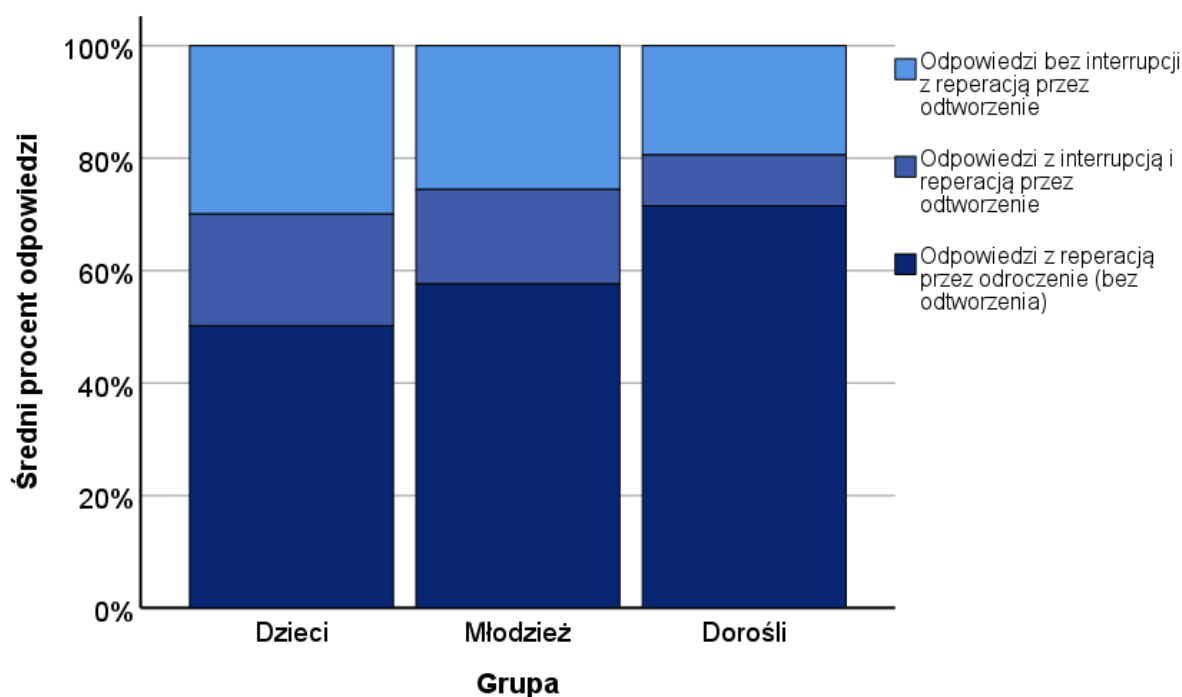
#### **6.4.2 Test hipotezy 1.3.2. Częstość stosowania reperacji ukrytych (strategia odraczania) wzrasta wraz z wiekiem.**

Aby zweryfikować hipotezę 1.3.2. na temat rosnącej częstości reperacji ukrytych (reperacje przez odroczenie) wraz z wiekiem postanowiono porównać częstości odpowiedzi z tego typu reperacjami w warunku błędnego dopełnienia w trzech badanych grupach wiekowych.

Rysunek 28 przedstawia częstość tego typu odpowiedzi na tle częstości odpowiedzi z reperacją (jawną) przez odtworzenie (bez interupcji oraz z przerwaniem wypowiedzi z reparamandum). Z przedstawionych danych wynika, że dzieci średnio w połowie odpowiedzi poprawiały powtarzane zdania poprzez reperację ukrytą wypowiedzi (powtarzane błędne zdanie od razu otrzymywało poprawne, tj. zgodne z obrazkiem dopełnienie). W starszych grupach wiekowych odpowiedzi takich było więcej. Z racji tego, że dane przedstawiają proporcje (odsetek odpowiedzi, w których dokonano reperacji ukrytej), czyli mieszczą się w zakresie 0-1 i nie spełniają założenia o normalności rozkładów w poszczególnych grupach, postanowiono zastosować nieparametryczny test Jonkheere-Terpstry, który jest właściwy dla danych porządkowych z założeniem o istnieniu trendu dla kolejnych populacji (grupa wiekowa).

## Rysunek 28

Średnie proporcje częstości odpowiedzi w warunkach błędnego dopełnienia: zestawienie dla trzech grup wiekowych



Wynik testu okazał się istotny statystycznie,  $T_{JT}(2, N = 149) = 4750,5$ ;  $z = 3,66$ ;  $p < 0,001$ ;  $\tau_b = 0,235$ , co oznacza że istnieje trend wzrostowy odsetka odpowiedzi z reperacją ukrytą (przez odroczenie) wraz z wiekiem. Na podstawie wielkości siły efektu można jednak uznać, że związek częstości reperacji ukrytych z wiekiem jest słaby. Porównania *post hoc* pomiędzy poszczególnymi grupami przy poziomie istotności skorygowanej metodą Bonferroniego dla wielu porównań wykazały, że w sposób istotny odsetek odpowiedzi z reperacją ukrytą odróżnia grupę dorosłych od grupy dzieci ( $p = 0,001$ ) oraz grupy młodzieży ( $p = 0,013$ ). Różnice między dziećmi i młodzieżą okazały się nieistotne statystycznie ( $p = 0,338$ ).

Na podstawie otrzymanych wyników można uznać, że wraz z wiekiem (od późnego dzieciństwa do wczesnej dorosłości) wzrasta częstość stosowania reperacji ukrytych przez odracanie, co **pozwała na uznanie hipotezy 1.3.2. jako częściowo potwierdzonej**. Jednocześnie niewielka siła efektu sugeruje, że są jeszcze inne ważne czynniki wpływające na częstość stosowania reperacji ukrytych w zadaniu *Powtarzanie zdań*.

### **6.4.2.1 Dodatkowa analiza (1): Stosowanie reperacji ukrytych w zadaniu *Powtarzania zdań* a tempo mówienia**

Ze względu na to, że moment przzerwania wypowiedzi w warunku „błédnego podmiotu” zwiázany był w znacznej mierze z tempem mówienia, postanowiono sprawdzić czy analizowana powyżej częstość stosowania reperacji ukrytych w warunku „błédnego dopełnienia” także wykazywać będzie takie związki.

Biorąc pod uwagę istotne różnice wynikające z przynależności do grupy wiekowej przeprowadzono analizę korelacji cząstkowej między tempem mówienia (w warunku neutralnym) i odsetkiem reperacji ukrytych (w warunku błédnego dopełnienia) przy kontroli zmiennej „wiek”. Współczynnik korelacji Pearsona wyniósł  $r = -0,643$ ;  $p < 0,001$ . Ujemna korelacja oznacza, że im szybsze było tempo mówienia badanej osoby, tym rzadziej stosowała reperacje ukryte<sup>170</sup>. Korelacja ta jest umiarkowana i wyjaśnia około 41% wariacji (z wyłączeniem wariacji wyjaśnianej przez wiek).

Na podstawie powyższych danych można stwierdzić, że osoby mówiące wolniej mają większą szansę na modyfikację buforowanej wypowiedzi, przez co łatwiej im dokonać reperacji ukrytej. Jest to spójne z obserwacją, że strategia odraczania wiąże się ze spowolnieniem, czy nawet chwilowym zawieszeniem artykulacji (Postma i Kolk, 1993).

### **6.4.2.2 Dodatkowa analiza (2): Spowolnienie tempa mówienia w wyniku stosowania strategii odraczania podczas reperacji ukrytych w różnych grupach wiekowych**

Mając na uwadze przesłankę, że strategia odraczania w reperowaniu błédów mowy wiąże się ze spowolnieniem lub nawet chwilowym przzerwaniem i zawieszeniem artykulacji (Postma i Kolk, 1993), a także wykazany powyżej zwiázek między tempem mówienia, a częstością reperacji przez odraczanie, postanowiono poszerzyć analizę o sprawdzenie, w jakim stopniu skuteczne stosowanie reperacji ukrytych wiązało się z obniżeniem tempa mówienia w każdej z grup wiekowych. W tym celu porównano średnie tempo mówienia w każdej z grup wiekowych w warunku neutralnym (zgodność zdania z treścią obrazka) ze średnim tempem mówienia w próbach z reperacją ukrytą w warunku „błédnego dopełnienia” (warunek błédnego dopełnienia traktowany był jako warunek eksperymentalny, natomiast warunek neutralny – jako warunek kontrolny). Porównań tempa mówienia dokonano za pomocą dwuczynnikowej analizy wariacji w schemacie mieszanym: 3 (grupa wiekowa) x 2 (warunek

<sup>170</sup> Ujemna korelacja sugeruje, że wolniejsze tempo mówienia sprzyjało stosowaniu reperacji ukrytych, co wpłynęło na wielkość różnic między grupami wiekowymi. Osoby młodsze, mówiące generalnie wolniej, miały większą szansę na wprowadzenie reperacji ukrytej, niż osoby starsze, których tempo mówienia jest wyższe. Tym bardziej zaobserwowane różnice między grupami wskazują, na znaczenie zmian rozwojowych postępujących wraz z wiekiem, dla zwiększenia częstości stosowania reperacji ukrytych.

zadania: w. neutralny vs w. błędnego dopełnienia). Czynnikiem mierzonym wewnątrz osób był warunek zadania eksperymentalnego, natomiast czynnikiem między osobami była przynależność do jednej z grup wiekowych. Zmienną zależną było tempo mówienia (średnia liczba sylab wypowiedzianych na sekundę). Analiza wariancji została przeprowadzona w modelu jednozmiennowym (*univariate repeated measures ANOVA*)<sup>171</sup>. Uzyskane zostały następujące efekty:

1. Uzyskano istotny efekt główny przynależności do grupy wiekowej  $F(2, 123) = 84,04$ ;  $p < 0,001$ ;  $\eta_p^2 = 0,577$ . Porównania *post hoc* za pomocą testu GT2 Hochberga wykazały istotne różnice między każdą z grup wiekowych. Oznacza to, że dzieci mówiły w najwolniejszym tempie, w porównaniu do młodzieży i dorosłych, natomiast dorośli mówili najszybciej w porównaniu z dwiema młodszymi grupami wiekowymi<sup>172</sup>.
2. Uzyskano istotny efekt główny warunku zadania (warunek neutralny vs warunek błędnego dopełnienia):  $F(1, 123) = 260,38$ ;  $p < 0,001$ ;  $\eta_p^2 = 0,679$ . Tempo mówienia w warunku błędnego dopełnienia było istotnie niższe niż w warunku neutralnym.
3. Uzyskano nieistotny efekt interakcji wieku oraz warunku zadania  $F(2,123) = 0,182$ ;  $p = 0,834$ ;  $\eta_p^2 = 0,003$ , co oznacza że różnice tempa między warunkami były zbliżone we wszystkich grupach wiekowych.

Okazuje się więc, że biorąc pod uwagę oba warunki zadania łącznie, dorośli mówią w najszybszym tempie spośród wszystkich grup wiekowych, natomiast tempo mówienia dzieci jest najwolniejsze. Jednocześnie wypowiedzenie poprawnego zdania, powstającego poprzez dokonanie reperacji ukrytej wiąże się z istotnym spowolnieniem tempa mówienia we wszystkich grupach wiekowych. Brak interakcji między czynnikami („wiek” oraz „warunek zadania”) pokazuje, że spadek tempa mówienia w wyniku realizacji reperacji ukrytej był zbliżony w każdej grupie wiekowej<sup>173</sup> i wynosił średnio 0,6 sylaby na sekundę (w grupie dzieci spadek wyniósł 0,61 syl./sek., w grupie młodzieży 0,58 syl./sek., a w grupie dorosłych

171 Rozkłady tempa mówienia w podgrupach wiekowych zbliżone są do rozkładu normalnego (wynik testu Shapiro-Wilka w żadnym z porównań nie uzyskał zadanego poziomu istotności  $\alpha=0,05$ ). Wykazują one także jednorodność wariancji międzygrupowej (test Levine’a także dał rezultat nieistotny statystycznie dla obu zmiennych w każdej z badanych podgrup), spełniając tym samym podstawowe założenia dla zastosowania testów parametrycznych. (Załącznik F)

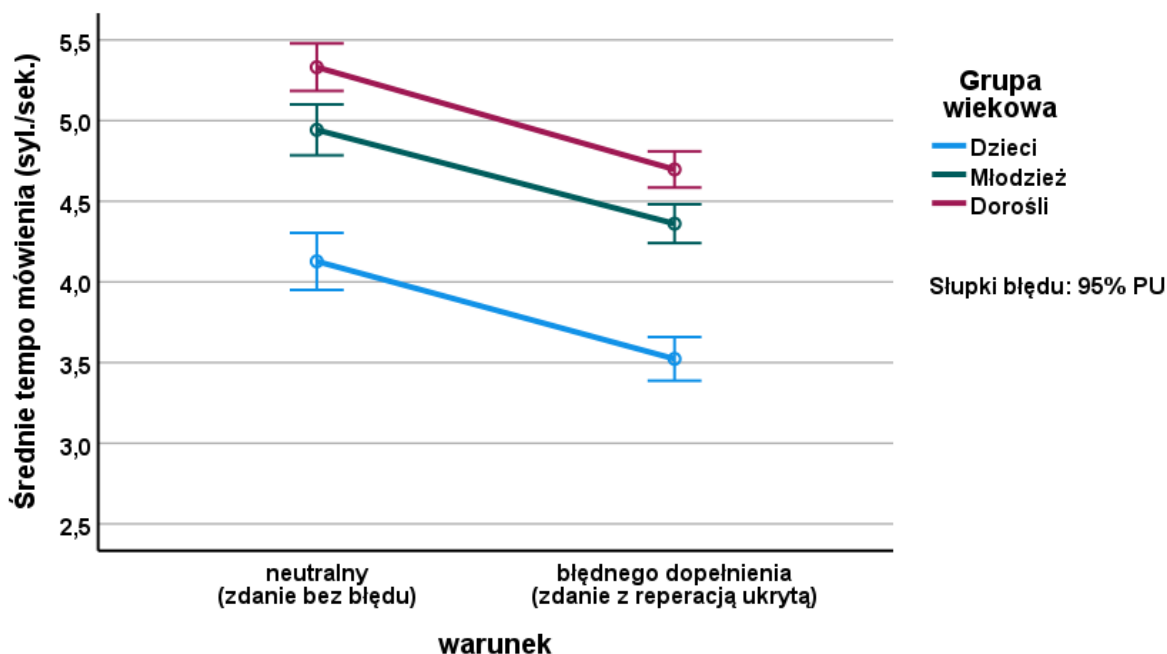
172 Wcześniejszy wynik porównań porównań międzygrupowych tempa mówienia przedstawiony w sekcji 6.4.1.2., w którym nie uzyskano istotnych różnic między grupą dorosłych i młodzieży dotyczył jedynie warunku neutralnego.

173 Analiza procentowego spadku tempa mówienia w wyniku zastosowania reperacji ukrytej w porównaniu do tempa mówienia w warunku neutralnym przy użyciu nieparametrycznego testu Jonkheere-Terpstry wykazała jednak istotne różnice między grupami wiekowymi:  $T_{JT}(2, N = 126) = 2193$ ;  $z = -1,91$ ;  $p = 0,031$  (istotność jednostronna);  $\tau_b = -0,132$ . Oznacza to, że w grupie dzieci procentowe opóźnienie tempa mówienia podczas stosowania strategii odraczania było największe (istotnie różne od opóźnienia zarejestrowanego u adolescentów i w grupie osób dorosłych). Efekt ten jest jednak marginalny (współczynnik determinacji  $R^2 = 0,017$ ).

0,63 syl./sek.). Zmiany tempa mówienia przy dokonywaniu reperacji ukrytych w porównaniu do tempa w warunkach neutralnych w poszczególnych grupach wiekowych przedstawia rysunek 29.

### Rysunek 29

Zestawienie różnic w tempie mówienia między zdaniem bez błędu a zdaniem z reperacją ukrytą



## 6.5 Kontrola procesu mówienia a poziom funkcji wykonawczych

Druga grupa hipotez odnosiła się do współzmienności funkcji wykonawczych i wybranych aspektów kontroli procesu mówienia. Aby zweryfikować sformułowane na ten temat hipotezy wykonano następujące działania.

W pierwszym kroku postanowiono sprawdzić, czy – zgodnie z przewidywaniami – poziom badanych funkcji wykonawczych jest zróżnicowany w badanej próbie i przyjmuje różne wartości w zależności od wieku (przynależności do danej grupy wiekowej) osób badanych. Statystyki opisowe wybranych do analizy wskaźników funkcji wykonawczej prezentuje tabela 30.

**Tabela 30***Statystyki opisowe dla wskaźników poziomu funkcji wykonawczych*

Grupa wiekowa	<i>M</i>	<i>SD</i>	<i>Me</i>	<i>SKE</i>	<i>K</i>
<i>Koszt konfliktu w zadaniu z flankerami (kontrola hamowania)</i>					
Dzieci ( <i>n</i> = 46)	25,5	17,2	25	0,749	1,258
Młodzież ( <i>n</i> = 50)	8,3	11,7	5	3,084	13,154
Dorośli ( <i>n</i> = 50)	8,2	7,1	7,5	2,259	7,332
<i>Procent poprawnych odpowiedzi w zadaniu 1-wstecz (odświeżanie pamięci roboczej)</i>					
Dzieci ( <i>n</i> = 39)	79,1	8,9	79,7	-0,110	-0,841
Młodzież ( <i>n</i> = 46)	89	5,8%	90,6	-1,043	0,974
Dorośli ( <i>n</i> = 50)	94,6	3,8	95,7	-0,921	0,256
<i>Procent błędów perseweracyjnych w teście sortowania kart (elastyczność poznawcza)</i>					
Dzieci ( <i>n</i> = 46)	20,9	7,3	18,8	0,635	0,071
Młodzież ( <i>n</i> = 49)	17,7	5,6	17,2	0,514	0,56
Dorośli ( <i>n</i> = 50)	13,3	4,4	12,5	0,948	1,067

*Adnotacja.* *n* – liczba obserwacji; *M* – średnia; *SD* – odchylenie standardowe; *Me* – mediana, *SKE* – skośność; *K* – kurtoza;

Z uwagi na to, że żaden z analizowanych rozkładów zmiennych nie wykazuje dostatecznego podobieństwa do rozkładu normalnego, do porównań międzygrupowych zastosowano nieparametryczny test *H* Kruskala-Wallisa, uzyskując następujące wyniki:

1. Porównanie wielkości kosztu konfliktu<sup>174</sup> w zadaniu z flankerami w poszczególnych grupach wiekowych dało wynik  $H(2, N = 146) = 42,82$ ;  $p < 0,001$ ;  $\eta^2_H = 0,285$ , który pozwala uznać, że między grupami występują istotne różnice pod względem poziomu funkcji kontroli hamowania. Porównania *post hoc* w poszczególnych parach przy poziomie istotności skorygowanej metodą Bonferroniego dla wielu porównań wykazały, że w sposób istotny wybrany wskaźnik funkcji hamowania odróżnia grupę dzieci od grupy młodzieży ( $p < 0,001$ ) oraz grupy dorosłych ( $p < 0,001$ ). Różnice w wynikach młodzieży i osób dorosłych okazały się nieistotne statystycznie ( $p > 0,05$ ).

<sup>174</sup> Różnica między poprawnością odpowiedzi w warunku z flankerami zgodnymi, a poprawnością odpowiedzi w warunku z flankerami niezgodnymi (zob. sekcja 5.4.2.2.3.)

2. Wynik porównania międzygrupowego dla poprawności wykonania zadania pamięciowego 1-wstecz okazał się istotny statystycznie:  $H(2, N = 135) = 69,28$ ;  $p < 0,001$ ;  $\eta^2_H = 0,51$ , co oznacza, że między grupami wiekowymi wystąpiły znaczące różnice pod względem odświeżania pamięci roboczej. Przeprowadzone następnie porównania *post hoc* z korektą Bonferroniego dla wielu porównań wykazały, że każda z grup wiekowych różniła się istotnie od pozostałych ( $p < 0,001$ ) pod względem poprawności wykonania zadania pamięciowego 1-wstecz. Poziom odświeżania pamięci roboczej był więc najwyższy u osób dorosłych, średni u młodzieży i najniższy u dzieci.
3. Zestawienie wyników testu sortowania kart (WCST) wykazało, że pod względem odsetka błędów perseweracyjnych wystąpiły istotne różnice między badanymi grupami wiekowymi:  $H(2, N = 145) = 34,21$ ;  $p < 0,001$ ;  $\eta^2_H = 0,227$ . Porównania *post hoc* przy zadanym poziomie istotności z korektą Bonferroniego dla porównań wielokrotnych wykazały, że w sposób istotny wybrany wskaźnik elastyczności poznawczej odróżnia grupę dorosłych od grupy młodzieży ( $p < 0,001$ ) oraz grupy dzieci ( $p < 0,001$ ). Różnice w wynikach dzieci i młodzieży okazały się nieistotne statystycznie ( $p = 0,212$ ).

Na podstawie uzyskanych wyników można uznać, że poziom testowanych funkcji wykonawczych w badanej próbie przejawia dynamikę rozwojową, która różnicuje uczestników badania w poszczególnych grupach wiekowych. Jednocześnie zmienność międzygrupowa przedstawia odmienny obraz zróżnicowania dla poszczególnych funkcji:

- a) poziom kontroli hamowania odróżnia grupę dzieci od pozostałych grup wiekowych, przyjmując wartość zbliżoną do poziomu prezentowanego przez dorosłych już w grupie młodzieży;
- b) funkcja odświeżania pamięci roboczej zmienia się wraz z wiekiem, osiągając coraz wyższe poziomy w kolejnych badanych grupach wiekowych<sup>175</sup>;
- c) poziom elastyczności poznawczej nieznacznie zmienia się na przestrzeni późnego dzieciństwa nie różnicując istotnie wyników dzieci oraz młodzieży, natomiast wyraźnie wzrasta w grupie osób dorosłych, odróżniając poziom wykonania zadania w tej grupie od wyniku osób w grupach młodszych.

Po ustaleniu charakterystyki testowanych funkcji wykonawczych w badanych grupach wiekowych, przystąpiono do testowania drugiej grupy hipotez badawczych, mówiących o

---

<sup>175</sup> Pod uwagę wzięto jedynie wynik zadania 1-wstecz, jednak obrazuje to również poziom wykonania zadania w warunku 2-wstecz, które okazało się za trudne dla większości dzieci oraz dużej części młodzieży (zob. sekcja 5.4.2.2.2.).

związkach między poziomem funkcji wykonawczych a poziomem kontroli procesu mówienia. Mając na względzie odstępstwa rozkładów zmiennych od rozkładu normalnego, do analiz statystycznych postanowiono wykorzystać nieparametryczny odpowiednik korelacji  $r$  Pearsona w postaci współczynnika  $\rho$  Spearmana (Conover, 1999; George i Mallery, 2019).

### **6.5.1 Test hipotezy 2.1.1. Odporność selekcji leksykalnej na dystrakcję semantyczną zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w okresie od późnego dzieciństwa przez adolescencję do wcześniej dorosłości.**

Aby zweryfikować hipotezę 2.1.1 mówiącą o wzroście odporności selekcji leksykalnej na dystrakcję wraz ze wzrostem poziomu funkcji wykonawczych zamierzano przeprowadzić analizę korelacji, sprawdzając czy odporność selekcji leksykalnej wykazuje związki ze sprawnością funkcji wykonawczych. Za wskaźnik odporności selekcji leksykalnej na dystrakcję przyjęto różnicę poprawności odpowiedzi między warunkiem neutralnym (bez dystrakcji), a warunkiem interferencji (z dystrakcją) w zadaniu nazywania obrazków. Przyjęto przy tym, że im większa różnica w poprawności nazywania obrazków tym mniejsza odporność selekcji leksykalnej na dystrakcję. W sekcji 6.2.2. (test hipotezy 1.1.2.) wykazano jednak, że osoby badane cechowały się wysoką odpornością na dystrakcję semantyczną i wprowadzenie do zadania dodatkowych bodźców interferujących z nazwami obrazków nie spowodowało istotnego pogorszenia poprawności odpowiedzi w grupie dorosłych i młodzieży, przyczyniając się jedynie do niewielkiego, ale istotnego spadku poprawności nazywania w grupie dzieci. Z tego względu nie jest możliwe zweryfikowanie hipotezy na podstawie wyników uzyskanych w całej przebadanej próbie i zostanie ona sprawdzona jedynie w odniesieniu do rezultatów uzyskanych przez dzieci.

W celu weryfikacji (w grupie dzieci) hipotezy 2.1.1. o wzroście odporności selekcji leksykalnej na dystrakcję wraz ze wzrostem poziomu funkcji wykonawczych przeprowadzono analizę korelacji zmiennej „odporność selekcji leksykalnej na dystrakcję” (której wskaźnikiem był spadek poprawności nazywania obrazków pod wpływem interferencji) z wynikami testów funkcji wykonawczych. Aby uniknąć artefaktów z analizy usunięto wyniki uczestników z grupy dzieci, u których wartość zmiennej „odporność selekcji leksykalnej” wyniosła poniżej 0 (przypadki, w których poprawność odpowiedzi w warunku bez dystrakcji była niższa niż w warunku interferencji). Uzyskane rezultaty przedstawia tabela 31.

Obliczone współczynniki korelacji Pearsona pomiędzy zmienną „odporność selekcji leksykalnej na dystrakcję” a zmiennymi „kontrola hamowania”, „odświeżanie pamięci roboczej” i „giętkość poznawcza” nie osiągnęły poziomu zadanej istotności statystycznej ( $\alpha =$

0,05), przez co nie można odrzucić hipotezy zerowej i stwierdzić istnienia między nimi związku monotonicznego. Otrzymane wyniki **nie pozwalają zatem na pozytywną weryfikację hipotezy 2.1.1.** i uznanie, że odporność selekcji leksykalnej (u dzieci) zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych.

**Tabela 31**

*Wartość współczynników korelacji między odpornością selekcji leksykalnej na dystrakcję a poziomem funkcji wykonawczych u dzieci.*

	Kontrola hamowania <i>n</i> = 34	Odświeżanie pamięci roboczej <i>n</i> = 30	Giętkość poznawcza <i>n</i> = 34
Odporność selekcji leksykalnej na dystrakcję	$\rho = 0,106$	$\rho = 0,078$	$\rho = 0,207$
	$p = 0,550$	$p = 0,683$	$p = 0,240$

*Adnotacja.* *n* – liczba obserwacji;  $\rho$  – współczynnik korelacji Spearmana; *p* – wartość *p* dla współczynnika korelacji Spearmana

### **6.5.2 Test hipotezy 2.1.2. Wielkość efektu interferencji semantycznej maleje wraz ze wzrostem poziomu funkcji wykonawczych u osób w okresie od późnego dzieciństwa przez adolescencję do wcześniej dorosłości.**

Aby zweryfikować hipotezę 2.1.2 i sprawdzić, czy wielkość efektu interferencji semantycznej maleje wraz ze wzrostem poziomu funkcji wykonawczych u osób w trzech wyróżnionych grupach wiekowych, przeprowadzono analizę korelacji zmiennej „wielkość efektu interferencji semantycznej” ze wskaźnikami funkcji wykonawczych przy kontroli zmiennej „wiek”. Obliczone współczynniki korelacji między zmiennymi przedstawia tabela 32. W żadnej z analizowanych par nie przekroczyły one zadanego poziomu istotności statystycznej ( $\alpha = 0,05$ ), przez co nie ma wystarczających podstaw do odrzucenia hipotezy zerowej. Otrzymane wyniki nie pozwalają zatem na pozytywną weryfikację hipotezy 2.1.2. i uznanie, że wielkość efektu interferencji maleje wraz ze wzrostem poziomu funkcji wykonawczych u osób w różnym wieku. **Hipoteza 2.1.2. nie została potwierdzona.**

**Tabela 32**

Wartość współczynników korelacji dla analizy związków między efektem interferencji semantycznej a poziomem funkcji wykonawczych

	Kontrola hamowania $n = 146$	Odświeżanie pamięci roboczej $n = 135$	Giętkość poznawcza $n = 145$
Wielkość efektu interferencji semantycznej	$\rho = -0,006$	$\rho = -0,020$	$\rho = 0,099$
	$p = 0,505$	$p = 0,816$	$p = 0,235$

Adnotacja. Wartości współczynnika korelacji przy kontroli zmiennej „wiek”;  $n$  – liczba obserwacji;  $\rho$  – współczynnik korelacji Spearmana;  $p$  – wartość  $p$  dla współczynnika korelacji Spearmana

### 6.5.2.1 Dodatkowa analiza: Nazywanie obrazków bez dystrakcji, a poziom funkcji wykonawczych

W związku z tym, że obie hipotezy szczegółowe 2.1.1. oraz 2.1.2. nie znalazły potwierdzenia w danych empirycznych postanowiono wykonać dodatkowe analizy, sprawdzające na ile uzyskane rezultaty zgadzają się z raportowanymi przez Shao i in. (2012) związkami między szybkością nazywania obrazków a funkcjami wykonawczymi. W cytowanym badaniu autorzy uznają, że, kontrola wykonawcza wykazuje związki z procesem mówienia, co przekłada się na szybkość nazywania obrazków. Raportowane przez nich współczynniki korelacji dotyczą więc szybkości nazywania obrazków, a nie efektu interferencji.

**Tabela 33**

Wartości współczynników korelacji uzyskane w badaniu Shao i in. (2012)

Funkcje wykonawcze (zastosowane zadania)	Czas reakcji (RT)	
	Nazywanie przedmiotów	Nazywanie czynności
Hamowanie ( <i>stop signal</i> )	0,45*	0,45*
Pamięć robocza ( <i>operation span</i> )	-0,54*	-0,38
Giętkość poznawcza ( <i>shape-colour</i> )	0,36	0,36

Adnotacja. \* $p < 0,05$ ; \*\* $p < 0,01$

**Tabela 34**

Wartość współczynników korelacji między wskaźnikami funkcji wykonawczych a poprawnością oraz szybkością nazywania obrazków w warunku bez interferencji (warunek neutralny)

Funkcje wykonawcze	<i>n</i>	Czas reakcji (RT)
Kontrola hamowania (zad. z flankerami)	133	0,292**
Odświeżanie p. roboczej (I-wstecz)	131	-0,240**
Giętkość poznawcza (WCST)	140	0,135

Adnotacja. Wartości współczynnika korelacji *r* Pearsona przy kontroli zmiennej „wiek”; \**p* < 0,05; \*\**p* < 0,01; *n* – liczebność próby

Analizując wartości współczynników korelacji przedstawione w tabelach 33 oraz 34<sup>176</sup> można dojść do wniosku, że w przeprowadzonym na potrzeby niniejszej rozprawy badaniu, kontrolując zmienną „wiek” uzyskano podobne zależności, jak cytowani Shao i in. (2012). W obu badaniach czas reakcji (nazywania obrazków) w warunku bez dystrakcji istotnie dodatnio koreluje z miarą kontroli hamowania i ujemnie koreluje z miarą odświeżania pamięci roboczej<sup>177</sup>. Siła wykazanego związku jest jednak niższa niż w raporcie Shao i in. (2012).

Zgodnie z powyższym, pomimo braku pozytywnej weryfikacji hipotez szczegółowych 2.1.1. oraz 2.1.2. **nie można odrzucić hipotezy ogólnej 2.1.**, mówiącej o tym, że poziom kontroli selekcji leksykalnej wiąże się z poziomem funkcji wykonawczych. Jak pokazała dodatkowa analiza, kontrola procesu mówienia (w tym kontrola selekcji leksykalnej) już na poziomie nazywania obrazków bez dystrakcji wykazuje związki z miarami funkcji wykonawczych. Szybkość nazywania koreluje dodatnio z miarą hamowania (im mniejszy koszt hamowania, tym krótszy czas nazywania) oraz ujemnie ze wskaźnikiem odświeżania pamięci roboczej (im niższa sprawność funkcji pamięci roboczej, tym dłuższy czas nazywania).

176 Zdecydowano się zastosować ten sam test statystyczny, który wykorzystany został przez Shao i in., (2012), tj. analizę korelacji *r* Pearsona. Aby możliwe było zastosowanie parametrycznego testu *r* Pearsona, usunięto w poszczególnych grupach wiekowych obserwacje zidentyfikowane jako nietypowe i odstające od średniej. Dzięki temu rozkłady zmiennych przyjęły kształt zbliżony do normalnego.

177 W oryginalnym raporcie Shao i in. (2012, s. 1935) wykonali dodatkowe obliczenia, uwzględniające specyfikę rozkładu wyników i wykazali, że przełączanie (*shifting*) może mieć znaczenie w większości prób nazywania obrazków przedstawiających czynności oraz tylko w sporadycznych, bardzo powolnych próbach nazywania obiektów.

### **6.5.3 Test hipotezy 2.2. Efektywność automonitorowania procesu mówienia wykazuje związek z poziomem funkcji wykonawczych.**

W celu weryfikacji hipotezy 2.2. przeprowadzono analizę korelacji cząstkowych między wskaźnikami sprawności automonitorowania procesu mówienia a miarami funkcji wykonawczych, przy kontroli zmiennej „wiek”, testując w ten sposób hipotezy szczegółowe 2.2.1. oraz 2.2.2.:

– Hipoteza 2.2.1. Dokładność monitorowania wewnętrznego zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w różnym wieku.

– Hipoteza 2.2.2. Szybkość monitorowania wewnętrznego zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w różnym wieku.

Uzyskane wyniki przedstawia tabela 35. Świadczą one o zróżnicowanym, ale pozytywnym związku sprawności automonitorowania procesu mówienia ze wszystkimi komponentami 3-czynnikowego modelu funkcji wykonawczych: kontroli hamowania, odświeżania pamięci roboczej oraz giętkości poznawczej.

1. Negatywny związek kontroli hamowania z dokładnością monitorowania wewnętrznego oznacza, że wraz ze wzrostem kontroli hamowania (mniejszy koszt konfliktu) wzrasta dokładność monitorowania.
2. Pozytywny związek odświeżania pamięci roboczej z dokładnością monitorowania oznacza, że wraz ze wzrostem poziomu tej funkcji wykonawczej, wzrasta także dokładność monitorowania. Wynik ten jest zgodny z przewidywaniami wysuniętymi na podstawie badania Levelta i Wheeldon (1995), w którym w efekcie obciążenia pamięci roboczej dodatkowym zadaniem odliczania nastąpił niewielki spadek poprawności rozpoznań monitorowanych fonemów oraz opóźnienie czasu reakcji. Nie udało się natomiast wykazać podobnego związku między poprawnością zadania 1-wstecz a szybkością monitorowania (nieistotna korelacja między miarą odświeżania a czasem efektywnego monitorowania)<sup>178</sup>.
3. Pozytywny związek giętkości poznawczej z szybkością monitorowania wewnętrznego oznacza, że wraz ze wzrostem poziomu giętkości (mniejszy odsetek błędów perseweracyjnych) skraca się czas przebiegu monitorowania wewnętrznego.

---

178 Kiedy jednak przeprowadzono dodatkową analizę korelacji wyłącznie w grupie osób dorosłych (właściwa grupa odniesienia pod względem wieku dla uczestników badania Levelta i Wheeldon, 1995) przy wykorzystaniu wskaźnika odświeżania pamięci roboczej w postaci „odsetka poprawnych odpowiedzi w wariacie zadania 2-wstecz”, uzyskano istotny wynik:  $r = -0,3$ ;  $p = 0,033$ . Sugeruje to, że w grupie osób dorosłych ( $n = 48$ , po usunięciu obserwacji nietypowych) wraz ze wzrostem poziomu funkcji odświeżania, odnotowuje się niewielki spadek czasu potrzebnego na efektywne monitorowanie wewnętrzne (wzrost szybkości monitorowania).

Z uzyskanych rezultatów wynika, że efektywność monitorowania wewnętrznego wykazuje istotne, choć słabe związki z poziomem funkcji wykonawczych. Dokładność i szybkość monitorowania w sposób zróżnicowany i ze stosunkowo niską siłą efektu korelują z miarami funkcji wykonawczych. Na podstawie tych przesłanek można uznać hipotezę 2.2. oraz szczegółowe hipotezy 2.2.1 oraz 2.2.2 za częściowo potwierdzone.

**Tabela 35**

*Wartość współczynników korelacji między sprawnością monitorowania wewnętrznego a poziomem funkcji wykonawczych*

	Kontrola hamowania	Odświeżanie pamięci roboczej	Giętkość poznawcza
Sprawność monitorowania wewnętrznego	$n = 145$	$n = 134$	$n = 144$
Dokładność monitorowania	-0,170*	0,383**	-0,154
Szybkość monitorowania	-0,058	-0,079	0,224**

*Adnotacja.* Wartości współczynnika korelacji  $\rho$  Spearmana przy kontroli zmiennej „wiek”;  $n$  – liczba obserwacji; \* $p < 0,05$ ; \*\* $p < 0,01$

#### **6.5.4 Test hipotezy 2.3. Sposób dokonywania reperacji wiąże się z poziomem funkcji wykonawczych**

W celu weryfikacji hipotezy 2.3. przeprowadzono analizę korelacji między zmiennymi dotyczącymi sposobu dokonywania reperacji a miarami funkcji wykonawczych, testując w ten sposób hipotezy szczegółowe 2.3.1. oraz 2.3.2.:

- Hipoteza 2.3.1. Szybkość dokonywania interupcji błędnych wypowiedzi zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w różnym wieku.
- Hipoteza 2.3.2. Częstość stosowania reperacji wewnętrznych przez strategię odraczania zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w różnym wieku.

W związku z wcześniejszym wykazaniem wpływu tempa mówienia na sposób dokonywania reperacji (zob. sekcja 6.4.2.1.), postanowiono włączyć tę zmienną – obok zmiennej „wiek” do modelu analizy korelacji, jako kowariant. Z uzyskanych wyników, które prezentuje tabela 36 wynika, że jedynie stosowanie reperacji ukrytych w warunku błędnego dopełnienia przy

kontroli wieku oraz tempa mówienia uczestników wykazuje istotne negatywne związki z poziomem kontroli hamowania (pomiar kosztu konfliktu w zadaniu z flankerami). Oznacza to, że im wyższy poziom kontroli hamowania prezentują badane osoby (coraz niższa wartość kosztu konfliktu), tym więcej reperacji ukrytych (przez odroczenie) są w stanie zastosować. W związku z tym, że rola funkcji wykonawczych w sposobie dokonywania reperacji w zadaniu *Powtarzanie zdań* okazała się niewielka, hipoteza szczegółowa 2.3.1. nie została potwierdzona, a hipotezę 2.3.2. należy uznać za jedynie częściowo potwierdzoną.

**Tabela 36**

*Wartość współczynników korelacji między dokonywaniem reperacji a poziomem funkcji wykonawczych*

	Kontrola hamowania	Odświeżanie pamięci roboczej	Giętkość poznawcza
Sposób dokonywania reperacji	$n = 146$	$n = 135$	$n = 144$
Szybkość dokonywania interupcji	0,064	-0,162	0,042
Stosowanie reperacji ukrytych	-0,166*	0,162	-0,122

*Adnotacja.* Wartości współczynnika korelacji  $\rho$  Spearmana przy kontroli zmiennych „wiek” i „tempo mówienia”;  $n$  – liczba obserwacji; \* $p < 0,05$ ; \*\* $p < 0,01$

## 7 Dyskusja i wnioski

Podstawowym celem niniejszej pracy jest analiza rozwojowych zmian w kontroli procesu mówienia oraz sprawdzenie na ile mechanizmy kontroli mowy wykazują związki z bardziej ogólnymi procesami kontroli poznawczej. Przedstawione w poprzednim podrozdziale wyniki przeprowadzonych badań były odpowiedzią na postawione pytania, odnoszące się do trzech wyróżnionych w rozdziale 5. aspektów kontroli mówienia obecnych w przejęzyczeniu i jego korekcie. Dyskusja tych wyników poprowadzona zostanie w niniejszej części według porządku weryfikacji hipotez.

### 7.1 Dyskusja wyników

#### 7.1.1 Kontrola selekcji leksykalnej

Zgodnie z przyjętym w niniejszej pracy modelem produkcji mowy Levelta (1989, 1998, 1999b, 2000; Levelt i in., 1999) leksykalne błędy mowy zachodzą w momencie zyskania przez niewłaściwą lemmę poziomu aktywacji, pozwalającego na kodowanie jej na kolejnych poziomach produkcji mowy. Źródłem tej aktywacji może być kontekst językowy, zarówno związany z wypowiedzią, jak i całkowicie wobec niej niezależny. W wykorzystanej w badaniu procedurze *Nazywania obrazków* kontekst ten podlegał manipulacji poprzez prezentowanie, tuż przed wyświetleniem obrazka, wyrazu semantycznie powiązanego z nazwą tego obrazka. Okazało się, że cała badana próba przedstawiła stosunkowo wysoki poziom odporności selekcji leksykalnej na dystrakcję, co owocowało średnią poprawnością odpowiedzi w warunku interferencji (dystrakcja obecna) na poziomie ponad 95%. Mimo tego, udało się zaobserwować różnice w kontroli selekcji leksykalnej wynikające z wieku, tym samym pozytywnie weryfikując hipotezę 1.1. mówiącą o tym, że poziom kontroli selekcji leksykalnej wzrasta wraz z wiekiem. Analiza szczegółowych hipotez pozwala dokładniej zobrazować poznawczo-rozwojowe aspekty kontroli leksykalnej.

Po pierwsze wykazano, że w kolejnych grupach wiekowych osiągnany jest coraz wyższy odsetek poprawnych odpowiedzi (poprawna nazwa obrazka), co wspiera hipotezę 1.1.1. mówiącą, że skuteczność selekcji leksykalnej wzrasta wraz z wiekiem. Jest to zgodne z wcześniejszymi badaniami, w których poprawność nazywania obiektów na obrazkach wzrastała w kolejnych grupach wiekowych, zarówno w zadaniach z prostym nazywaniem obrazków (Rosinski, 1977) jak i w zadaniach wymagających opisu obrazka (Hanley i in., 2016).

W przypadku raportowanego badania w warunku nazywania obrazków bez dystrakcji średnia poprawność odpowiedzi u dzieci wyniosła ponad 96%, a u młodzieży i dorosłych ponad 99%. Różnice w poprawności między dziećmi i dwoma pozostałymi starszymi grupami wiekowymi były istotne statystycznie, natomiast nie wykazano różnic między młodzieżą i dorosłymi. Może to oznaczać, że dojrzała (typowa dla dorosłych) kontrola selekcji leksykalnej osiągnięta jest już w adolescencji (Shao i in., 2013). Alternatywne wyjaśnienie tego zjawiska związane jest z czułością zastosowanej metody na potencjalne różnice między adolescentami i dorosłymi. Niewykluczone, że przy zastosowaniu innych wariantów tego zadania, np. nazywanie na czas (Vitkovitch i in., 1993) lub z kontrolą czynników lingwistycznych, takich jak frekwencja, wiek akwizycji, prawdopodobieństwo fonotaktyczne itp., możliwe byłoby uzyskanie bardziej zróżnicowanego poziomu wykonania w grupie adolescentów i młodych dorosłych (por. Newman i German, 2005). Na podstawie uzyskanych wyników nie można jednak stwierdzić różnic między młodzieżą i dorosłymi, co oznacza częściowe potwierdzenie hipotezy 1.1.1.

Zgodnie z przewidywaniami zastosowanie zewnętrznej dystrakcji semantycznej znalazło swoje odzwierciedlenie w pogorszeniu wyników (spadek odsetka poprawnych odpowiedzi), ale jedynie w grupie dzieci. Choć nieznaczne spadki poprawności odpowiedzi odnotowano także w grupach młodzieży i dorosłych, okazały się one nieistotne statystycznie. Można zatem uznać, że wyniki te tylko częściowo wspierają hipotezę 1.1.2., a skuteczność selekcji leksykalnej pod wpływem zewnętrznej dystrakcji semantycznej obniża się jedynie u dzieci. Jednocześnie, zestawiając wyniki w obu warunkach zadania można uznać, że odporność selekcji leksykalnej na zewnętrzną dystrakcję semantyczną wzrasta wraz z wiekiem. Potwierdza to porównanie międzygrupowe, które wykazało istniejący trend rozwojowy w zmniejszającej się wraz z wiekiem różnicy między poprawnością odpowiedzi w warunku neutralnym a warunku interferencji. Oznacza to, że wraz z wiekiem dystrakcja semantyczna w coraz mniejszym stopniu wpływa na poprawność selekcji leksykalnej. Należy przy tym zaznaczyć, że efekt ten okazał się niewielki, a różnice między „sąsiadującymi” grupami wiekowymi zatarte (istotna różnica wystąpiła jedynie między grupą dzieci i dorosłych). Dodatkowa analiza pokazała również, że wpływ dystrakcji semantycznej na poprawność odpowiedzi u dzieci polegał głównie na blokowaniu nazywania obrazków, nie powodując istotnego zwiększenia odsetka błędów leksykalnych. O ile więc dystrakcja taka, pojawiając się w sposób spontaniczny i nieoczekiwany może wpłynąć na przebieg procesu mówienia (Clark, 1996), skutkując przejęzyczeniem (Harley, 1984; 2001), o tyle podawana w sposób systematyczny i jawny w paradygmacie *picture-word interference*, działa mobilizująco

na procesy kontrolne, nie dopuszczając do realizacji konkurencyjnie wzbudzonych jednostek leksykalnych. Może mieć też znaczenie to, że osoby badane mogły zapoznać się wcześniej z obrazkami, które będą nazywać. Kawachi (2002) wykazał, że mowa przygotowana wcześniej zawiera mniej leksykalnych błędów paradygmatycznych niż mowa spontaniczna. Według niego, podobny efekt wprawy występuje także w badaniach eksperymentalnych, zmniejszając proporcje błędów względem mowy spontanicznej. O ile więc kontrola selekcji leksykalnej jest niższa u dzieci i wykazują one mniejszą odporność na dystrakcję semantyczną niż osoby starsze, nie musi się to przekładać na większy odsetek popełnianych przez nie przejęzyczeń. Jest to zgodne z danymi obserwacyjnymi pokazującymi, że dzieci nie popełniają więcej przejęzyczeń leksykalnych niż dorośli (Jaeger, 2005; Wijnen, 1992) (choć rzadziej je poprawiają), natomiast okresowo pojawiają się u nich rozwojowe zaburzenia płynności mowy (Evans, 1985; Levelt, 1998; Sabin i in., 1979; Tarkowski, 1992), co w kontekście uzyskanych wyników zostanie skomentowane w dalszej części tej sekcji.

Innym badanym skutkiem zewnętrznej dystrakcji semantycznej w nazywaniu obrazków jest spowolnienie czasu reakcji (czasu nazywania) w warunku interferencji względem warunku neutralnego. Efekt ten, nazywany efektem interferencji semantycznej udało się wykazać w odniesieniu do wszystkich grup wiekowych. Tym samym przy udziale użytkowników języka polskiego zreplikowano znany z badań nad użytkownikami innych języków efekt opóźnienia czasu reakcji pod wpływem dystrakcji semantycznej (Glaser i Döngelhoff, 1984; Schriefers i in., 1990; por. Kujalowicz i Zajdler, 2009). W związku z tym, że dystrakcja semantyczna w zastosowanej procedurze oddziałuje najsilniej na proces selekcji leksykalnej (Levelt, 1993) wyniki te można interpretować zgodnie z hipotezą 1.1.5., według której szybkość selekcji leksykalnej zmniejsza się pod wpływem zewnętrznej dystrakcji semantycznej u osób w różnym wieku: zarówno dzieci, młodzieży, jak i dorosłych.

Choć efekt interferencji wystąpił w całej badanej próbie, to jego siła (wielkość opóźnienia) uzależniona była od przynależności do danej grupy wiekowej, wykazując różnice rozwojowe związane z wiekiem. Po pierwsze czas potrzebny na wygenerowanie nazwy obrazka (niezależnie od warunku) był najdłuższy w grupie dzieci i skracał się w kolejnych grupach wiekowych. Wykazano więc, zgodnie z hipotezą 1.1.4., że szybkość selekcji leksykalnej skraca się wraz z wiekiem. Po drugie, opóźnienie nazywania wynikające z zewnętrznej dystrakcji semantycznej także było największe u dzieci i malało w kolejnych grupach wiekowych, potwierdzając słuszność przypuszczenia, że wielkość efektu interferencji semantycznej maleje wraz z wiekiem (hipoteza 1.1.6.). Jest to zgodne z wcześniejszymi wynikami uzyskanymi przez Rosinskiego (1977), w których ogólna wielkość interferencji

była większa w przypadku młodszych osób. Jednocześnie uzyskane przez badacza różnice wynikające z wieku były na tyle małe, że na ich podstawie nie można było stwierdzić jakościowo odmiennych sposobów semantycznego przetwarzania języka w produkcji mowy (nazywaniu obrazków). Podobnie, nowsze badanie (Jerger i in., 2002) pokazało, że efekt interferencji semantycznej w nazywaniu obrazków (przy SOA = -150 ms) jest większy u dzieci 5-7 letnich niż u adolescentów, ale pomimo różnic w szybkości selekcji leksykalnej, przebieg procesu produkcji mowy jest podobny u dzieci, młodzieży i dorosłych. Analogiczne wnioski na podstawie analizy błędów popełnianych w zadaniu nazywania obrazków wyciągnęli Budd i in., (2011), twierdząc że dotychczasowe badania dowodzą istnienia raczej różnic ilościowych niż jakościowych w organizacji systemu produkcji mowy u dzieci i osób dorosłych. Według tych badaczy obserwowane różnice zacierają się wraz z wiekiem i postępującym rozwojem powiązań między poziomami sieci leksykalnej.

W podobnym kluczu można interpretować wyniki niniejszego badania (zob. też wyniki zadania *Monitorowanie sylab*), przy czym należy odróżnić ilościowe zmiany rozwojowe związane ze wzrostem szybkości selekcji leksykalnej oraz zwiększającą się wraz z wiekiem odpornością selekcji leksykalnej na dystrakcję semantyczną od efektów jakościowych, uzewnętrzniających się w płynności mowy. Inaczej mówiąc, na podstawie uzyskanych wyników można wskazać, że ilościowe różnice w zakresie przebiegu procesów planowania i programowania mowy, mają swoje przełożenie na jakość zewnętrznie obserwowanego procesu mówienia u osób w badanych grupach wiekowych. Zostanie to dokładniej omówione poniżej.

Jak już wspomniano, w pomiarze czasów reakcji osób badanych ze wszystkich grup wiekowych ujawnił się efekt interferencji semantycznej, przy czym u dzieci był on najsilniejszy i stopniowo „słabł” wraz z wiekiem, osiągając najniższą wartość w grupie dorosłych. Skalę opóźnienia w dostępie leksykalnym, wywoływanego przez dystrakcję semantyczną pomogła zobrazować dodatkowa analiza, polegające na przeliczeniu bezwzględnych wartości czasu reakcji na procentowe wskaźniki opóźnienia nazywania (porównanie o ile procent wolniejsza była reakcja w warunku interferencji względem warunku neutralnego). Średnia różnica opóźnienia między kolejnymi grupami wiekowymi wyniosła około 9 punktów procentowych. O ile więc w przypadku dorosłych uczestników badania opóźnienie nazywania wyniosło średnio 23%, to u młodzieży przyjmowało ono wartość 32%, a w przypadku dzieci wzrastało do poziomu 41%. Tak duże opóźnienie nie powinno pozostać bez wpływu na płynność mówienia, dlatego przeprowadzono dodatkową

analizę dystrybucji pauz wypełnionych<sup>179</sup> w obu warunkach zadania *Nazywanie obrazków*. Ujawniła ona, że w warunku interferencji istotnie wzrastał (względem warunku neutralnego) odsetek pauz wypełnionych. Zjawisko to było szczególnie zauważalne u dzieci i młodzieży, gdyż w przypadku osób dorosłych różnice między warunkami nie uzyskały poziomu zadanej istotności statystycznej. Biorąc pod uwagę poprawność odpowiedzi, częstość pauz wypełnionych oraz efekt opóźnienia nazywania zauważyć można interesujący trend rozwojowy związany z jakościowymi różnicami międzygrupowymi, jakie powstały pod wpływem zewnętrznej dystrybucji semantycznej. Przedstawia je tabela 37. Zauważyć można, że zewnętrzna dystrybucja semantyczna najsilniej ingeruje w proces produkcji mowy u dzieci, nie tylko znacznie opóźniając selekcję leksykalną, ale też wpływając na zwiększenie pauz wypełnionych oraz obniżenie poprawności odpowiedzi (wzrost braków odpowiedzi). W grupie młodzieży efekt interferencji jest nie tylko mniejszy niż u dzieci, ale nie powoduje spadków poprawności odpowiedzi, choć nadal pogarsza płynność mowy, zwiększając liczbę pauz wypełnionych. Ostatnia z badanych grup wiekowych, młodzi dorośli, najlepiej radzi sobie z dystrybucją semantyczną, która w tej grupie wiekowej powoduje najmniejsze opóźnienia w nazywaniu oraz nie zwiększa liczby pauz wypełnionych, błędów leksykalnych ani braków odpowiedzi.

**Tabela 37**

*Efekty wpływu zewnętrznej dystrybucji semantycznej na wykonanie zadania Nazywanie obrazków: porównanie trzech grup wiekowych*

	Opóźnienie nazywania	Wzrost liczby pauz wypełnionych	Spadek poprawności odpowiedzi
Dzieci	x	x	x
Młodzież	x	x	—
Dorośli	x	—	—

*Źródło:* opracowanie własne

<sup>179</sup> Pauzy wypełnione były jedynym rejestrowanym dodatkowo zjawiskiem towarzyszącym nazywaniu obrazków, które można było powiązać z płynnością mowy. Nie rejestrowano powtórzeń, falstartów ani słów przerwanych w trakcie artykulacji. Jak wykazali Hartsuiker i Notebaert (2010) pauzy wypełnione są lepszym wskaźnikiem trudności w dostępie leksykalnym niż powtórzenia dźwięków, sylab, czy wyrazów.

Podsumowując, dostarczone przez zadanie *Nazywanie obrazków* wyniki pokazują ilościowe zmiany w zakresie kontroli selekcji leksykalnej, dokonujące się w okresie od późnego dzieciństwa, przez adolescencję, do wczesniej dorosłości. Są one spójne z wnioskami płynącymi z dotychczasowych badań nad produkcją mowy w nazywaniu obrazków u osób w różnym wieku i posługujących się innymi językami niż język polski. Należy jednak poczynić w tym miejscu ważne zastrzeżenie dotyczące zastosowanej w niniejszej pracy metody i jej związku z badanymi zmiennymi. Uzyskane efekty czasowe nazywania obrazków interpretowane są w kategoriach czasu potrzebnego na efektywną selekcję leksykalną. Wynika to z założenia, że użyta w procedurze dystrakcja semantyczna oddziałuje właśnie najsilniej na procesy leksykalne w produkcji mowy, skutkując opóźnieniem nazywania (Levelt, 1993). Czas reakcji mierzony w zadaniu nazywania obrazków (w obu warunkach) zawiera w sobie jednak cały zakres procesów związanych z produkcją mowy (produkcji wyrazu), a także poprzedzającą go identyfikacją wizualną obrazka. O ile więc różnica w czasie nazywania między warunkiem interferencji, a warunkiem neutralnym może być interpretowana jako efekt wydłużenia czasu efektywnej selekcji leksykalnej (spowolnienie dostępu leksykalnego w wyniku dystrakcji), to już różnice w czasie nazywania występujące między grupami wiekowymi mogą wynikać z różnic w tempie przebiegu procesów planowania i programowania wypowiedzi na innych poziomach niż selekcja leksykalna. Zastrzeżenie to dotyczy np. potwierdzenia hipotezy 1.1.4., mówiącej o tym, że szybkość selekcji leksykalnej wzrasta się wraz z wiekiem. Choć zostało wykazane, że czas reakcji (czas nazywania) skraca się w kolejnych grupach wiekowych, nie można w całości tego efektu przypisać zmianom rozwojowym w tempie selekcji leksykalnej, gdyż na czas nazywania składają się jeszcze inne – wspomniane wcześniej – poziomy produkcji mowy. Nie można tym samym wykluczyć, że dłuższy czas nazywania w młodszych grupach wiekowych wiąże się nie tylko z dłuższą selekcją leksykalną, ale także z tym, że więcej czasu zajmuje młodszemu osobom identyfikacja obrazka, czy generowanie planu fonetycznego. Z tego względu różnic w czasie nazywania nie można utożsamiać wprost z dokładną miarą różnic w czasie w selekcji leksykalnej. Niestety metoda nazywania obrazków nie pozwala na precyzyjne wydzielenie odcinków czasowych z rejestrowanego czasu reakcji, aby odróżnić czas trwania poszczególnych elementarnych procesów. Jest to mankament dotyczący różnych metod chronometrycznych, np. decyzji leksykalnej (*lexical decision task*) (Levelt, 1993), gdzie pomija się wprowadzić późniejsze etapy programowania mowy, ale nadal czas reakcji zawiera w sobie – oprócz czasu potrzebnego na efektywną selekcję leksykalną – czas identyfikowania obiektu na obrazku oraz czas reakcji motorycznej (wybór klawisza). Pewnym

rozwiązaniem jest zastosowanie metody neuroobrazowania, która – dzięki wykazaniu zróżnicowanej aktywności kory mózgowej w poszczególnych etapach produkcji mowy – pozwala sekwencjonować składowe procesy zaangażowane w nazywanie obrazków (Indefrey, 2011; Indefrey i Levelt, 2004) Na przykład Piai i in. (2014) obserwowali przy użyciu megnetoencefalografii aktywność mózgu osób badanych podczas zadania nazywania obrazków z towarzyszeniem różnych dystraktorów. Wykazali oni, że współzawodnictwo między konkurencyjnymi jednostkami leksykalnymi odzwierciedla się w większej aktywności neuronów w lewym górnym zakręcie czołowym (SFG – *left superior frontal gyrus*) w przedziale 350-650 ms od rozpoczęcia nazywania (od wyświetlenia obrazka). Aktywność ta była największa w przypadku dystrakcji o charakterze semantycznym w porównaniu do warunku dystrakcji niepowiązanej oraz dystraktorów tożsamy z nazwą obrazka. Co więcej, różnice w aktywności SFG odzwierciedlone były także w standardowym pomiarze czasów reakcji (opóźnienie nazywania wywołane dystrakcją). Przyszłe badania nad selekcją leksykalną u różnych grup badanych (np. różnice rozwojowe wynikające z wieku lub różnice między grupami klinicznymi) powinny uwzględnić metodologię bardziej czułą na czasowe przebiegi poszczególnych etapów planowania i programowania mowy, aby precyzyjnie oddzielić efekty specyficzne dla danego poziomu przetwarzania języka.

### **7.1.2 Automonitorowanie wewnętrzne**

Monitorowanie własnej mowy pozwala na weryfikację generowanego tekstu i sprawdzenie go pod kątem zgodności z intencją komunikacyjną, regułami poprawności językowej, czy dopasowania do (domniemanej) wiedzy, uwagi i możliwości odbioru interlokutora (Clark, 1996; Levelt, 1989; Postma, 2000). Szczególnie interesujący w kontekście przeprowadzonych badań jest aspekt sprawdzania wypowiedzi pod względem jej zgodności z intencją, gdyż pozwala to na wykrycie potencjalnego przejęzyczenia tuż przed jego wypowiedzeniem (*wewnętrzna pętla monitorowania*) lub już po jego wyartykułowaniu (*zewnętrzna pętla monitorowania*). Ze względu na to, że monitorowanie elementów subleksykalnych (fonemów, sylab) w produkowanej mowie jest ekwiwalentne wobec monitorowania błędów mowy – zarówno pod względem zgodności z teorią podwójnej pętli monitorowania (Levelt i in., 1999; Wheeldon i Levelt, 1995; Schiller, 2005) jak i na podstawie danych empirycznych (Nooteboom i Quené, 2019) – do sprawdzenia rozwojowych różnic w zakresie efektywności monitorowania wewnętrznego wykorzystano zadanie *Monitorowania sylab*.

Uzyskane w zadaniu wyniki pokazują, że pomiędzy badanymi grupami wiekowymi widać progres w zakresie poprawności oraz szybkości wykonania zadania, co wspiera

hipotezę 1.2., że automonitorowanie (wewnętrzne) zwiększa swoją efektywność wraz z wiekiem. Szczegółowe analizy pokazały interesujące zjawiska, które zostaną omówione poniżej w porządku podyktowanym kolejnością testowanych hipotez: w pierwszej kolejności podsumowane zostaną wnioski z weryfikacji hipotez o rozwojowych zmianach w monitorowaniu wynikających z wieku osób badanych, a w kolejnym kroku przedstawione zostaną wnioski z weryfikacji hipotez o strukturalnym (językowym) uwarunkowaniu procesu monitorowania.

Poprawność wykonania zadania, polegająca na właściwym rozpoznaniu sylaby docelowej w strukturze nazw prezentowanych obrazków (lub stwierdzeniu jej nieobecności w danej nazwie) wzrastała w kolejnych badanych grupach wiekowych. Pozwala to na stwierdzenie, że monitorowanie wewnętrzne (zadanie wykonywane było bezgłośnie) zwiększa swoją dokładność wraz z wiekiem. Tym samym hipotezę 1.2.1. można uznać za potwierdzoną.

Podobnie, szybkość udzielania odpowiedzi w zadaniu zmieniała się wraz z wiekiem osób badanych. Dotyczyło to zarówno warunku zgodności, w którym poprawna odpowiedź wiązała się z pozytywnym rozpoznaniem sylaby docelowej w strukturze monitorowanego wyrazu, jak i w warunku niezgodności, w którym poprawna odpowiedź wymagała negatywnego stwierdzenia o braku obecności sylaby docelowej w strukturze wyrazu. Generalnie więc, udało się wykazać różnice między wszystkimi grupami wiekowymi w uzyskiwanych czasach reakcji, co pozwala zgodnie z hipotezą 1.2.2. na stwierdzenie, że szybkość monitorowania wewnętrznego skraca się wraz z wiekiem. Tym samym uzyskane wyniki korespondują z wcześniejszymi, nielicznymi badaniami nad wzrastającą wraz z wiekiem efektywnością monitorowania mowy (Sasisekaran i Weber-Fox; 2012; por. Hanley i in., 2016).

Należy też zauważyć, że podobnie, jak w zadaniu *Nazywanie obrazków*, nie odnotowano różnic jakościowych pomiędzy tym, jak materiał bodźcowy przetwarzany był przez osoby badane z różnych grup wiekowych. W związku z tym, że zadanie wymagało wykrywania sylab na różnych pozycjach wyrazu oraz w wyrazach różnej długości, możliwe było nie tylko porównanie średniego czasu monitorowania między grupami, ale także wykreślenie profili monitorowania sylab na poszczególnych pozycjach wyrazów i w wyrazach różnej długości. W żadnym z warunków odpowiedzi dzieci, młodzieży i dorosłych nie odbiegały od siebie pod względem innym niż ilościowy. Kierunek zależności był natomiast zbliżony w każdej z grup wiekowych. Pozwala to na stwierdzenie tożsamości procesu monitorowania u badanych osób, znajdujących się w wybranych okresach

rozwojowych. Jednocześnie, różnice w monitorowaniu wynikające ze struktury przetwarzanego tekstu (opisane poniżej) najsilniej uwypukliły się w grupie dzieci, co pokazuje, jak formalne aspekty języka stopniowo opanowywane są przez kształtujący się system produkcji mowy.

Kolejne postawione hipotezy dotyczyły tego, jakie znaczenie dla efektywności monitorowania mowy ma struktura wyrazu i miejsce zajmowane w niej przez sylabę docelową. Zgodnie z modelem Levelta i in. (1999; Hartsuiker, 2007; Hartsuiker i Kolk, 2001; Roelofs, 2005a; Schiller, 2005), w którym proces kodowania fonologicznego przebiega przyrostowo (*incremental*), również monitorowanie sekwencyjnie sprawdza kolejne elementy generowanej wypowiedzi. Założono więc, że wyrazy krótsze monitorowane będą szybciej niż wyrazy dłuższe oraz, że sylaby na początkowych pozycjach będą wykrywane w krótszym czasie niż sylaby na dalszych pozycjach przetwarzanego wyrazu. Otrzymane wyniki generalnie potwierdziły te przypuszczenia i obraz wykazanych zależności – o czym już wspomniano – był bardzo podobny w każdej z badanych grup wiekowych. Szczegółowe kwestie wymagają jednak dodatkowego komentarza, przy czym najpierw omówione zostaną wnioski dotyczące weryfikacji hipotezy 1.2.4, a następnie – hipotezy 1.2.3.

Zgodnie z hipotezą 1.2.4. wykazano, że czas efektywnego monitorowania sylab zależał od długości przetwarzanego wyrazu: monitorowanie wyrazów dwusylabowych zajmowało istotnie mniej czasu niż monitorowanie wyrazów trójsylabowych. Wpływ długości wyrazu na szybkość monitorowania był jednak uzależniony od wieku osób badanych, gdyż najsilniej ujawnił się w grupie dzieci, natomiast nie był istotny statystycznie w grupie osób dorosłych. Gdyby więc badanie prowadzone było wyłącznie wśród osób dorosłych – jak to ma miejsce w przypadku większości badań psycholingwistycznych nad produkcją mowy – otrzymane dane byłyby niejednoznaczne. Z jednej strony w grupie dorosłych wyniki ujawniły szybsze monitorowanie wyrazów krótszych, względem dłuższych, a z drugiej strony różnica ta okazała się na tyle mała, że nie przeszła testu istotności i nie może być podstawą do wnioskowania o istnieniu różnic. Dane takie nie dałyby więc rozstrzygnąć czy monitorowane wyrazy rzeczywiście przetwarzane były w sposób analityczny i przyrostowy, czy globalny. Mając jednak dane od osób w różnych grupach wiekowych, można dostrzec dynamikę rozwojową w zakresie szybkości monitorowania wewnętrznego. W przypadku dzieci oraz młodzieży czas monitorowania wydłuża się wraz ze wzrostem długości wyrazu, co potwierdza że struktura monitorowanego wyrazu sprawdzana jest „po kolei”, segment po segmencie (Wheeldon i Levelt, 1995; Schiller, 2005). W przypadku osób dorosłych wzrost czasu reakcji jest minimalny, co może wskazywać, że efekt ten ma marginalne znaczenie dla

efektywności automonitorowania produkcji mowy przez osoby dorosłe, przy niewielkich różnicach w długości wyrazu (różnica jednej sylaby).

Analogicznie hipoteza 1.2.3. zakładała, że w związku z przyrostowym charakterem procesu monitorowania, czas efektywnego monitorowania wewnętrznego (wskaźnik szybkości monitorowania) będzie wydłużał się wraz z zajmowaniem przez wykrywany element kolejnych pozycji w strukturze wyrazu. Przewidywanie to sprawdziło się w wynikach dostarczonych przez badanych ze wszystkich grup wiekowych, w odniesieniu do zestawienia pozycji pierwszej (nagłosowej) z pozostałymi pozycjami (ponownie w wynikach dzieci różnice między pozycjami były największe). Inaczej mówiąc, sylaby docelowe znajdujące się w nagłosie wykrywane były szybciej niż sylaby docelowe występujące w śródgłosie i w wygłosie. Wbrew przewidywaniom nie było istotnych statystycznie różnic w szybkości wykrywania sylab w śródgłosie i wygłosie. Co więcej, zauważalny był raczej niewielki trend o przeciwnym do zakładanego kierunku: sylaby docelowe w wygłosie średnio wykrywane były w minimalnie krótszym czasie niż sylaby docelowe w śródgłosie, przy czym różnice te w żadnej z badanych grup wiekowych nie przekroczyły poziomu zadanej istotności statystycznej. Efekt ten stanowi wyzwanie dla przyjętego modelu mówienia i monitorowania mowy, gdyż pokazuje załamanie przyrostowego charakteru monitorowania na środkowej sylabie. Poniżej przedstawiono możliwe interpretacje.

Po pierwsze, należy zauważyć, że „przyspieszenie” monitorowania po przekroczeniu punktu wyjątkowości wyrazu (*perceptual uniqueness point*) jest zjawiskiem raportowanym we wcześniejszych badaniach (Özdemir i in., 2007) i świadczy na rzecz teorii o odbiorczym charakterze pętli monitorowania mowy (Roelofs, 2020a, 2020b). Oznacza to, że od tego momentu, kiedy monitorowany wyraz nie dzieli wspólnych fonemów z innymi, niespokrewnionymi wyrazami, jego rozpoznanie staje się łatwiejsze, a efektywne monitorowanie wymaga mniej czasu. Materiał wykorzystany w niniejszym badaniu na potrzeby zadania *Monitorowanie sylab* nie był kontrolowany pod względem ilości segmentów poprzedzających w każdym wyrazie jego punkt wyjątkowości, przez co nie jest możliwe porównanie pod względem tego parametru użytych wyrazów z sylabą docelową w śródgłosie oraz wyrazów z sylabą docelową w wygłosie. Pobieźna analiza pozwala jednak stwierdzić, że dla większości z tych wyrazów (wyrazy 3-sylabowe) miejsce, w którym nie dzielą one już fonemów z innymi wyrazami w systemie występuje po drugiej sylabie (np. papuga) lub w ramach trzeciej sylaby (np. latawiec, kanapa). Po osiągnięciu tego miejsca przez monitor rozpoznawanie danego wyrazu powinno być łatwiejsze, przyspieszając tym samym proces monitorowania i podejmowania decyzji o obecności (lub nie) sylaby docelowej. Jeśli punkt

rozpoznawalności w wyrazie 3-sylabowym wypadałby po 2 sylabie, to tempo rozpoznawania sylaby w wygłosie mogłoby być wyższe niż sylaby w śródgłosie, skutkując tylko niewielkim przyrostem czasu potrzebnego na efektywne monitorowanie ostatniej sylaby względem sylaby na drugiej pozycji. Mogłoby to tłumaczyć, dlaczego w uzyskanych wynikach między wykryciem sylaby na drugiej pozycji, a wykryciem sylaby na trzeciej pozycji nie ma istotnych różnic, podczas gdy różnice takie obserwowane są między wykryciem sylaby na pierwszej i kolejnych pozycjach.

Powyższe wyjaśnienie ma jednak pewną wadę. W przypadku gdyby za brak różnic między czasem wykrycia sylaby w śródgłosie i nagłosie odpowiadało zjawisko punktu wyjątkowości percepcyjnej wyrazu, nadal powinna być obserwowana tendencja zwykła (wzrost ilości czasu potrzebnego na efektywne monitorowanie ostatniej sylaby względem sylaby środkowej), tymczasem wyniki w każdej z grup wiekowych pokazują raczej minimalny trend spadkowy – wykrycie sylaby w wygłosie zajmowało średnio mniej czasu niż wykrycie sylaby w śródgłosie. Wyjaśnienie tego wymaga uwzględnienia dodatkowych przesłanek.

Warto w tym kontekście spojrzeć na zadanie *Monitorowania sylab* pod kątem tego, jakie wymagania stawia przed osobami badanymi w zakresie świadomości językowej i metajęzykowej. Zdolność monitorowania swojej mowy i sprawdzania jej poprawności oraz dokonywanie w niej reperacji zaliczane bywa do wczesnych umiejętności metapoznawczych i świadomości językowej (Clark, 1978) lub wprost do kompetencji metajęzykowych (Levy, 1999). Kwarciak (1995) nazywa ten wczesny rodzaj refleksji nad językiem świadomością metajęzykową percepcyjną, wskazując na jej podstawowy charakter, oparty na analizie akustycznych cech strumienia mowy. Wynika ona z naturalnego zainteresowania substancją języka obecną już u małych dzieci, bazuje więc na wrodzonych mechanizmach organizujących percepcję słuchową. Pozostałe dwa typy wyróżnionej przez tego autora świadomości metajęzykowej, to świadomość naturalna (komunikacyjna) oraz analityczna. Obie bazują na doświadczeniu i jakiejś formie edukacji językowej, choć niekoniecznie sformalizowanej. Ta pierwsza (naturalna) ma charakter pre- lub ateoretyczny i służy głównie sprawnej komunikacji, druga (analityczna) natomiast wykorzystuje pewną usystematyzowaną teorię języka, przekazywaną chociażby w procesie nauki czytania i pisania. Kwarciak (1995) uważa, że spontaniczne korekty błędów składniowych oraz podział słowa na sylaby w trakcie nauki czytania i pisania należą do jakościowo różnych czynności, wymagających innych operacji na strukturze języka. Z tego względu można przyjąć, że automatyczne korekty spontanicznych przejęzyczeń, obserwowane już na bardzo wczesnych etapach ontogenezy

mowy (Clark, 1978; Laakso, 2010) oparte są na percepcyjnej świadomości metajęzykowej. Chociaż istnieją przesłanki, że paradygmat monitorowania elementów subleksykalnych języka wprowadzony przez Wheeldon i Levelta (1995) angażuje procesy kontrolne wykorzystywane w typowym monitorowaniu mowy służącym wykrywaniu błędów (Nooteboom i Quené, 2019), nie można wykluczyć, że wymaga także innego, zdobywanego w toku edukacji językowej, poziomu świadomości metajęzykowej.

Do zadania *Monitorowanie sylab* celowo wybrano właśnie sylaby, jako monitorowane elementy subleksykalne, ze względu na uczestniczące w badaniu dzieci, będące dopiero na początku formalnej edukacji metajęzykowej (okres wczesnoszkolny). Cytowane już wcześniej badanie Sasisekaran i Weber-Fox (2012) pokazało, że różne elementy strukturalne języka mogą być monitorowane z różną skutecznością, gdyż dla dzieci w wieku późnego dzieciństwa monitorowanie pojedynczych fonemów było trudniejsze niż monitorowanie rymu (monitorowanie fonemów zajęło więcej czasu). Analiza sylabowa pojawia się w doświadczeniu wcześniej niż analiza głoskowa (np. w postaci wyliczanek), często jeszcze poza kontekstem nauki czytania i pisania (Krasowicz-Kupis, 2004). Przygotowując zadanie do niniejszego badania założono więc, że monitorowanie sylab będzie wyzwaniem wymagającym umiejętności metajęzykowych adekwatnych dla poziomu rozwoju świadomości językowej najmłodszej grupy osób badanych. Być może jednak, w związku z tym, że sylaby w wyrazie mają przypisany akcent, rozpoznawanie ich na różnych pozycjach wyrazu wiąże się z dodatkowymi operacjami metafonologicznymi, wykraczającymi poza sam proces monitorowania procesu mówienia. Interesująco w tym kontekście prezentują się badania nad świadomością sylabową, w których proszono osoby badane o usuwanie poszczególnych sylab z powtarzanych przez nich wyrazów. Na ich podstawie Rosner i Simon (1971, za: Krasowicz-Kupis, 2004) ustalili, że usunięcie ostatniego segmentu słowa jest łatwiejsze niż usunięcie pierwszego (radziły sobie z tym wyzwaniem dzieci w młodszym wieku), a z największym trudem przychodziło badanym usuwanie środkowej sylaby. Trudno w tym miejscu rozciągać analogię między zadaniem usuwania sylab (konieczna jest analiza wyrazu oraz jego ponowna synteza z pominięciem środkowego segmentu), a zadaniem ich wykrywania w strukturze wyrazu, ale nie można wykluczyć, że przetworzenie sylab akcentowanych<sup>180</sup> także angażuje monitor w większym stopniu niż przetwarzanie sylab nieakcentowanych, zwiększając czas potrzebny na decyzję o obecności sylaby. Kwarciak cytuje badania wedle których sylaby akcentowane są dokładniej wymawiane przez dzieci i stają się często przedmiotem ich szczególnej uwagi (Echols i Newport, 1991, za: Kwarciak,

---

180 Dla języka polskiego typowy jest akcent na drugą sylabę od końca, a zatem na sylabę w śródgłosie w zastosowanych w zadaniu *Monitorowanie sylab* wyrazach 3-sylabowych.

1995). Możliwe więc, że dla rozpoznania sylaby docelowej w śródgłosie wyrazu 3-sylabowego, to że jest to sylaba akcentowana stanowi dodatkowy aspekt do przetworzenia, wydłużając czas monitorowania. Niestety nie można tego rozstrzygnąć na podstawie dostępnych danych. Wiadomo, że akcent może podlegać monitorowaniu i że – podobnie jak monitorowanie fonemów (Wheeldon i Levelt, 1995) – wykrywanie akcentu na kolejnych pozycjach w wyrazach 2- i 3-sylabowych odbywa się z coraz większym, przyrostowym opóźnieniem (Schiller i in., 2006). Poza tym Schiller (2005) na materiale języka niderlandzkiego wykazał jeszcze, że w wyrazach dwusylabowych największe przyrosty latencji w monitorowaniu dotyczą fonemów występujących w segmentach przed akcentowaną sylabą. Trudno więc stwierdzić, czy przypisanie do sylaby (niezależnie od jej pozycji w wyrazie) akcentu wydłuża czas jej efektywnego monitorowania. Kolejne badania powinny uwzględnić ten problem, aby przynieść jednoznaczne rozstrzygnięcia.

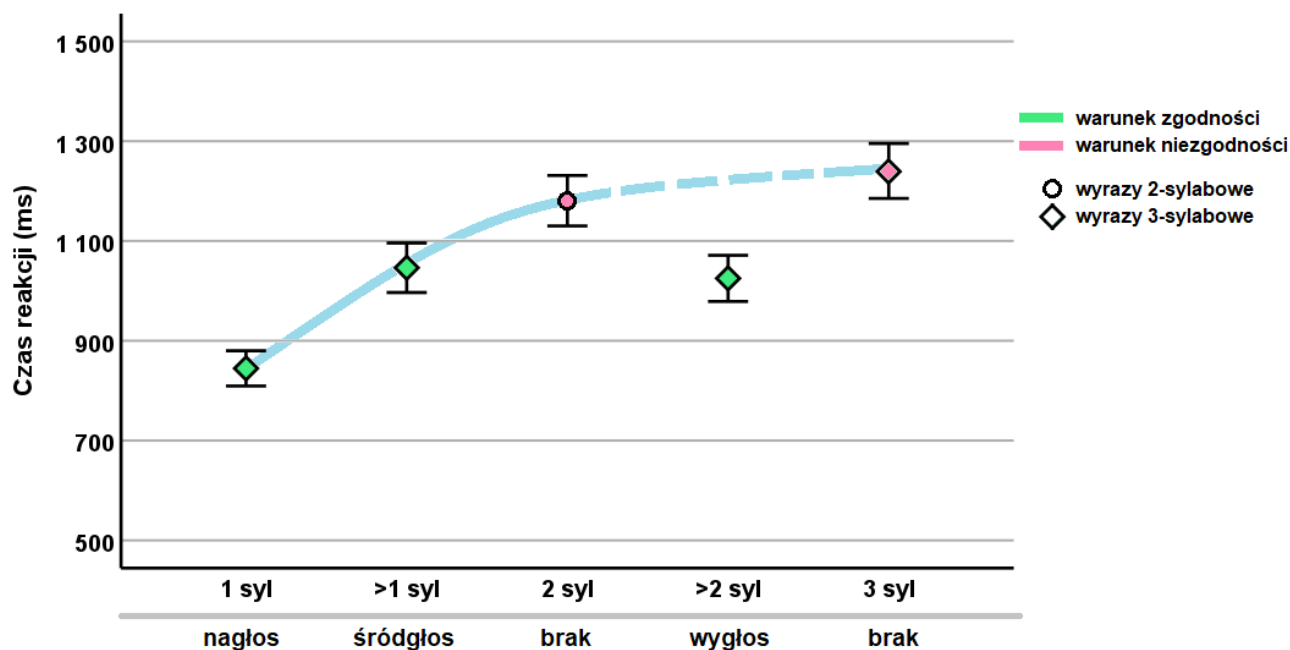
Możliwe jest zresztą jeszcze inne wyjaśnienie otrzymanych wyników, podkreślające nie tyle szczególne opóźnienia w przetwarzaniu sylab docelowych w śródgłosie (sylaba akcentowana), co wyjątkowo krótki czas monitorowania sylab w wygłosie. Aby to zobrazować, na rysunku 30 zestawiono średnie wartości czasów reakcji w obu warunkach zadania *Monitorowanie sylab* (warunek zgodności i warunek niezgodności) w zależności od tego, ile kolejnych sylab należało przetworzyć, aby udzielić poprawnej odpowiedzi. Wynika z niego, że wraz ze zwiększaniem się materiału językowego koniecznego do przetworzenia przed podjęciem decyzji o obecności lub nie sylaby docelowej w podanym wyrazie (nazwa obrazka), wydłużał się także czas reakcji. Tylko jeden z pięciu pomiarów przedstawia odstępstwo od wynikającego z modelu Levelta i in. (1999) przyrostowego charakteru kodowania fonologicznego i monitorowania jego postępów przez wewnętrzną pętlę percepcyjną. Pomiar czwarty<sup>181</sup> wyraźnie odstaje od pozostałych pomiarów. Możliwe więc, że krótszy czas reakcji na sylabę docelową w wygłosie niż w śródgłosie jest artefaktem, odzwierciedlającym bardziej jakąś anomalię pomiarową wynikającą z niedoskonałości zastosowanej metody, niż czasową charakterystykę procesu automonitorowania mowy.

---

181 Pomiar czasu reakcji dla wykrycia sylaby docelowej w wygłosie. Wykrycie sylaby docelowej na trzeciej pozycji wymagało przetworzenia dwóch pierwszych sylab w całości i przynajmniej jednego fonemu z segmentów składających się na ostatnią sylabę. Z tego względu pomiar ten został umieszczony pomiędzy monitorowaniem wyrazów dwusylabowych i trzysylabowych niezawierających sylaby docelowej (aby uznać, że w nazwie obrazka nie ma sylaby docelowej, konieczne jest przetworzenie całych wyrazów, a więc dwu pełnych sylab w trzecim pomiarze oraz trzech pełnych sylab w piątym pomiarze).

## Rysunek 30

Średnie wartości czasów reakcji we wszystkich pomiarach (oba warunki) zadania  
*Monitorowanie sylab*



Jak wspomniano wcześniej, materiał językowy użyty w zadaniu *Monitorowanie sylab* nie był kontrolowany pod względem ilości segmentów poprzedzających w każdym wyrazie jego punkt wyjątkowości. Podobnie było z innymi parametrami, które mogą mieć znaczenie dla szybkości przetwarzania mowy jak frekwencja użytych wyrazów (nazw obrazków) oraz sylab docelowych, prawdopodobieństwo fonotaktyczne, czy gęstość sąsiedztwa (*neighborhood density*). Wynika to ze sposobu doboru materiału do badania. Głównym kryterium było znalezienie wśród nazw obrazków z bazy IPNP (Szekely i in., 2004) wyrazów 3-sylabowych, które zawierałyby w swojej strukturze przynajmniej dwie sylaby otwarte CV, mogące być przedmiotem monitorowania (sylaby docelowe w nagłosie, śródgłosie i wygłosie wyrazów). Ograniczyło to znacznie możliwość wyboru materiału językowego przez co nie sformułowano dodatkowych kryteriów włączania i wyłączenia wyrazów z przygotowanego zestawu. Wiązało się to z ryzykiem, że materiał językowy nie będzie w pełni zbalansowany. Niestety podczas projektowania i przeprowadzania badania nie było jeszcze dostępnych aktualnych norm dla nazywania obrazków w języku polskim (por. Wolna i in., 2023), które pozwoliłyby sprawdzić parametry opracowywanego materiału. Nie można zatem wykluczyć,

że wyrazy użyte w którymś z pomiarów różnią się istotnie pod względem szybkości ich przetwarzania ze względu na jakiś nieuwzględniony parametr lingwistyczny.

Aby rzucić nieco światła na ten problem postanowiono sprawdzić częstość wystąpień w języku polskim wyrazów użytych w warunku zgodności jako nośników sylab docelowych w śródgłosie i wygłosie (pomiar 3. i 4. na rysunku 30) oraz wyrazów 3-sylabowych z warunku niezgodności (pomiar 5. na rysunku 30). W tym celu wykorzystano wyszukiwarke korpusową PELCRA, której zadano wyszukanie w NKJP wyrazów użytych w badaniu (pomiar 3-5). Zapytanie dotyczyło częstości wystąpień poszczególnych rzeczowników (nazw obrazków) w formie mianownikowej liczby pojedynczej. Uzyskane wyniki przedstawia tabela 38. Wynika z niej, że choć część wyrazów będących nośnikiem sylaby docelowej na 2. i 3. pozycji pokryła się, to generalnie wyrazy z sylabą docelową w wygłosie mają średnią częstość występowania w korpusie wyższą niż pozostałe zestawy.

**Tabela 38**

*Zestawienie częstości wystąpień w NKJP wyrazów użytych w warunku zgodności (śródgłosowa i wygłosowa pozycja sylaby docelowej) i niezgodności (wyrazy 3-sylabowe)*

Warunek zgodności		Warunek niezgodności			
Śródgłos		Wygłos		3 sylaby	
<i>M (SE)</i>		<i>M (SE)</i>		<i>M (SE)</i>	
2958,4 (977,7)		6384,9 (2985,3)		1670,1 (346,3)	
<i>Me</i>	1680	<i>Me</i>	3770	<i>Me</i>	1575
cebula	8646	korona	29309	koszula	3159
kapelusz	6192	cebula	8646	żarówka	1745
parasol	4882	pudełko	5064	widelec	1601
papuga	2081	butelka	4507	pomidor	1549
latawiec	1680	motyl(e)	3770	drabina	1438
kanapa	1347	rakieta	3348	patelnia	529
korale	1240	kanapa	1347		
komoda	325	korale	1240		
skarpeta	233	skarpeta	233		

*Źródło:* opracowanie własne na podstawie wyniku wyszukiwania podanych wyrazów przez wyszukiwarke PELCRA

Nie jest to precyzyjna miara częstości występowania danych leksemów w języku (por. *Słownik frekwencyjny polszczyzny współczesnej*, Kurcz i in., 1990), ale wystarczająco dobrze pokazuje, że między użytymi zestawami wyrazów mogą istnieć różnice wpływające globalnie na szybkość ich przetwarzania (niezależnie od miejsca sylaby docelowej). Może to ostatecznie rzutować na uzyskane w badaniu wyniki dotyczące monitorowania sylab. Dotyczy to zresztą nie tylko częstości wyrazów, ale także częstości sylab (lub poszczególnych segmentów), gdyż, jak wykazali Cholin, Levelt i Schiller (2006) frekwencja sylab (a przynajmniej sylaby nagłosowej) ma znaczenie dla szybkości monitorowania w produkcji mowy. Dalsze badania powinny uwzględnić ten problem i wykorzystać dostępne dane dotyczące statystyki języka polskiego, aby lepiej zrównoważyć materiał językowy służący weryfikacji hipotez psycholingwistycznych.

### **7.1.3 Reperacja błędów mowy**

Zauważalnym efektem procesu monitorowania jest korygowanie błędów we własnej mowie. Zewnętrzna pętla monitorowania (dominujący słuchowy kanał sprzężenia zwrotnego) pozwala odbierać własną mowę tak, jak mowę innych osób (Huettig i Hartsuiker, 2010; Levelt, 1989; Postma, 2000) i reagować na już wyartykułowany tekst. Wykrycie w nim błędu skutkuje przerwaniem danej wypowiedzi i wprowadzeniem do niej reperacji, najczęściej poprzez odtworzenie części wypowiedzi zawierającej błąd. Oprócz automonitorowania zewnętrznego postuluje się także istnienie wewnętrznej pętli monitorowania (Blackmer i Mitton, 1991; Levelt, 1989; Postma i Oomen, 2005; Hartsuiker i in., 2005), która mogłaby sprawdzać poprawność generowanej wypowiedzi jeszcze w jej postaci przeartykulacyjnej. Odpowiednio wczesne wykrycie w niej błędu miałyby skutkować przerywaniem wypowiedzi często jeszcze przed pełnym wypowiedzeniem błędnego słowa lub nawet dokonanie tzw. reperacji ukrytej, czyli korekty w planie wypowiedzi przed jej rzeczywistym wyartykułowaniem. Pytania badawcze zaadresowano do działania obu pętli monitorowania, zastanawiając się jak w okresie między późnym dzieciństwem a wczesną dorosłością zmienia się zdolność do korygowania własnych wypowiedzi. Pierwsza z postawionych hipotez szczegółowych (hipoteza 1.3.1.) przewidywała, że szybkość dokonywania interupcji w błędnej wypowiedzi będzie wzrastać wraz z wiekiem, odzwierciedlając rozwojowy progres w zakresie mobilizowania procesów naprawczych uruchamianych w wyniku wykrycia błędu przez zewnętrzną pętlę monitorowania. Druga hipoteza (hipoteza 1.3.2.) dotyczyła predykcji, że wraz z wiekiem – dzięki wzrostowi sprawności wewnętrznej pętli monitorowania – będzie wzrastać częstość stosowania reperacji ukrytych (pod uwagę wzięto reperacje ukryte

dokonywane przez odraczenie, por. Postma i Kolk, 1993). Do zweryfikowania hipotez wykorzystano zadanie *Powtarzania zdań*, będące modyfikacją paradygmatu wywoływania reperacji (*repair elicitation paradigm*) zaproponowanego przez van Wijka i Kempena (1987; por. Boland i in., 2005). Zadanie składało się z warunku kontrolnego (warunek neutralny: powtarzanie zdań bez błędu), warunku błędnego podmiotu (powtarzanie zdania z błędnym podmiotem) oraz warunku błędnego dopełnienia (powtarzanie zdania z błędnym dopełnieniem).

Do przetestowania hipotezy 1.3.1. posłużyły wyniki warunku błędnego podmiotu. Miarą szybkości dokonywania interupcji była długość wypowiedzi z reperandum (mierzona w sylabach), wyznaczająca odległość od początku zdania do momentu wprowadzenia interupcji (Levelt, 1983; 1989). Okazało się, że wbrew oczekiwaniom moment interupcji u badanych osób był bardzo zbliżony i we wszystkich grupach wiekowych wyniósł średnio między 4, a 5 sylabą od początku zdania. Nie wykazano zatem różnic między badanymi grupami wiekowymi pod względem szybkości dokonywania interupcji i nie potwierdzono hipotezy 1.3.1.

Aby lepiej zrozumieć powody uzyskania takich wyników przeprowadzono dodatkowe analizy. Po pierwsze wykazano, że chociaż nie ma różnic w szybkości dokonywania interupcji między badanymi grupami wiekowymi, to istnieje niewielki, ale istotny trend rozwojowy w zakresie częstości przerywania błędnych wypowiedzi (warunek błędnego podmiotu umożliwiał reperację błędu poprzez przerwanie reparandum i jego wznowienie lub wprowadzenie korekty po wypowiedzeniu pełnego błędnego zdania bez przerywania go). Inaczej mówiąc, wraz z wiekiem osoby badane częściej przerywały błędne wypowiedzi, nie doprowadzając reparandum do końca. W tym kontekście właściwe zinterpretowanie uzyskanych danych, wymagało uwzględnienia dodatkowych zmiennych. Biorąc pod uwagę przesłanki o wzrastającym wraz z wiekiem tempie mówienia (Sabin i in., 1979), postanowiono sprawdzić, czy zmienna ta mogła mieć wpływ na uzyskane wyniki. W pierwszym kroku potwierdzono, na podstawie materiału z warunku neutralnego (kontrolnego), że tempo mówienia (lub tempo wypowiedzi, por. Cholewiak, 2016) różnicowało osoby podzielone ze względu na przynależność do wyróżnionych grup wiekowych. Tempo mówienia dzieci było najniższe, podczas gdy tempo mówienia dorosłych było najwyższe, a czynnik wieku tłumaczył 37% wariacji w tempie mówienia. Podobnie, tempo mówienia okazało się wykazywać związki z szybkością przerywania wypowiedzi, co potwierdził ujemny wynik korelacji Pearsona (im szybsze tempo, tym później przerywane zdanie w warunku błędnego podmiotu). Na podstawie tych przesłanek, postanowiono w

dalszym kroku wykorzystać zmienną „tempo mówienia”, jako kowariant dla zmiennych „wiek<sup>182</sup>” i „szybkość interupcji”, których relacje były weryfikowane w związku z hipotezą 1.3.1. Zgodnie z ze sformułowanymi roboczo przewidywaniami, kontrola tempa mówienia przyniosła interesujące rezultaty. Otóż, korelacja cząstkowa przy kontroli zmiennej „tempo mówienia” pokazała istotne ujemne związki między wiekiem, a szybkością dokonywania interupcji. Okazało się więc, że przy stałej wartości tempa mówienia, szybkość dokonywania interupcji rzeczywiście wzrastała wraz z wiekiem (wraz z wiekiem skracala się długość wypowiedzi z reparandum). Tym samym hipoteza 1.3.1. otrzymała pośrednie wsparcie.

Chcąc wyjaśnić w jaki sposób tempo mówienia odegrało tak istotną rolę w zadaniu *Powtarzanie zdań*, należy odwołać się do zastosowanej procedury (zob. 5.4.2.1.3.). Została ona zaprojektowana w taki sposób, aby obrazek służący weryfikacji poprawności wypowiedzanego zdania pojawiał się na ekranie już w trakcie artykulacji. Z tego względu wyświetlenie obrazka powiązane było z klawiszem głosowym, który po zarejestrowaniu głosu osoby badanej (rozpoczęcie głośnego powtarzania zdania) uruchamiał wyświetlenie obrazka. Od wykrycia dźwięku mowy do wyświetlenia obrazka mijało 150 ms. Przy potencjalnych różnicach w tempie mówienia, ten krótki odcinek czasu mógł nieznacznie faworyzować osoby mówiące wolniej. Inaczej mówiąc, osoby mówiące w szybszym tempie potencjalnie wypowiadały więcej tekstu zanim dostały szansę na zweryfikowanie poprawności wypowiedzanego zdania.

Aby uniknąć artefaktów, przyszłe badania wykorzystujące paradygmat wywoływania reperacji (*repair elicitation paradigm*), powinny uwzględnić rolę tempa mówienia w wykonaniu zadań tego typu. Pewnym rozwiązaniem mogłoby być dynamiczne, uzależnione od tempa mówienia regulowanie czasu asynchronii między bodźcami. Osoby mówiące szybciej, mogłyby wcześniej otrzymywać wskazówkę dla weryfikacji swojej wypowiedzi, niż osoby mówiące w niższym tempie. Takie rozwiązania stosuje się np. w zadaniach mierzących kontrolę hamowania (np. Woźniak-Prus i Gambin, 2011).

Innym rozwiązaniem jest wykorzystanie bezwzględnego czasu trwania wypowiedzi z reparandum (*start to-cut-off* lub *error to cut-off*), jako miary szybkości przerywania wypowiedzi. Zmierzenie odcinka czasu, jaki upłynął od rozpoczęcia artykulacji (lub od wystąpienia błędu) jest alternatywnym sposobem wyznaczania czasu w jakim monitorowanie wykrywa błąd i uruchamia procesy naprawcze. W niniejszym badaniu nie wykonywano tego typu pomiarów, natomiast w ramach dodatkowej analizy obliczono potencjalny średni czas trwania wypowiedzi dla każdej z badanych osób na podstawie tempa mówienia (w warunku

---

182 Zastosowano tu zmienną ilorazową „wiek w miesiącach”, a nie trójwartościową zmienną porządkową „grupa wiekowa”

neutralnym) oraz liczby sylab wypowiedzianych do momentu wprowadzenia korekty (w warunku błędnego podmiotu). Zgodnie z przewidywaniami, ta nowa miara szybkości dokonywania interupcji okazała się różnicująca dla osób w różnym wieku, choć siła efektu była bardzo niska.

Ostatecznie więc uzyskane wyniki nie są rozstrzygające dla postawionej hipotezy. Można jednak na ich podstawie z pewną dozą ostrożności stwierdzić, że szybkość z jaką pętla monitorowania zewnętrznego uruchamia procesy naprawcze w wyniku wykrycia błędów wzrasta wraz z wiekiem, choć efekt ten może nie być widoczny ze względu na zmieniające się także z wiekiem ogólne tempo mówienia. Być może najważniejsze zmiany rozwojowe w tym zakresie zachodzą we wczesnym dzieciństwie, kiedy dzieci przechodzą od wykrywania i naprawiania błędów w oparciu o korekty zewnętrzne inicjowane przez inne osoby (*other-initiated repairs*) do interupcji wypowiedzi w tej samej turze konwersacyjnej, w oparciu o samodzielnie wykryty błąd (*self-initiated repairs*) (Clark, 1978; Clark, 1996; Forrester, 2008; Laakso, 2010; MacDonald i in., 2012). Późniejsze zmiany mogą być mniej uchwytnie, szczególnie ze względu na różnice w tempie mówienia oraz inne zakłócenia płynności mowy, wynikające z rozwijającej się kompetencji językowej.

W związku z powyższym należy uznać, że szybkość przerywania wypowiedzi nie jest istotnym wskaźnikiem rozwoju w zakresie monitorowania i reperowania mowy w okresie między późnym dzieciństwem a wczesną dorosłością. Należy też wziąć pod uwagę, że zastosowana procedura eksperymentalna stworzyła specyficzne warunki, w których szybkie dokonywanie interupcji było preferowaną formą reagowania. W toku codziennych rozmów opóźnienie korekty, czy nawet jej pominięcie mogą być bardziej pożądaną formą reakcji na błąd, jeśli przedłożenie płynności ponad poprawność miałyby uzasadnienie (Blackmer i Mitton, 1991; Seyfeddinipur i in., 2008).

W tym kontekście interesujące są wyniki testu drugiej z postawionych hipotez o rozwojowych zmianach w dokonywaniu reperacji. Reperacje ukryte w strategii odraczania mają tę szczególną właściwość, że ich zastosowanie pozwala pogodzić problem przetargu między poprawnością a płynnością. Inaczej mówiąc, jeśli błąd uda się wykryć odpowiednio wcześnie, to możliwe jest jego szybkie, przedartykulacyjne skorygowanie, które tylko w niewielkim stopniu zakłóci płynność mówienia (Hartsuiker, 2007). Tym samym możliwe jest utrzymanie poprawności wypowiedzi bez istotnej utraty jej płynności.

Według Hartsuikera i Kolka (2001) wyrazy znajdujące się na końcu buforowanego zdania są najbardziej predestynowane do tego, aby podlegać korektom wewnętrznym (ukrytym). Z tego względu do weryfikacji hipotezy 1.3.2. wykorzystano wyniki warunku

błédnego dopełnienia, w którym błédem obarczony był ostatni wyraz w powtarzanym zdaniu. Ze względu na to, że obrazek służący weryfikacji poprawności zdania wyświetlany był już 150 ms od rozpoczęcia mówienia, informacja o tym, czy dopełnienie zdania jest poprawne, czy nie, mogła dotrzeć do osoby badanej jeszcze przed wyartykułowaniem dopełnienia. Tym samym, jeśli dopełnienie nie zgadzało się z przedmiotem przedstawionym na obrazku, możliwe było wykrycie tego i rozpoczęcie korekty wypowiedzi jeszcze na poziomie planu przedartykulacyjnego. Ze względu na to, że działanie pętli wewnętrznej wymaga większego wysiłku niż działanie pętli zewnętrznej (Hartsuiker, Kolk i Martensen, 2005), przyjęto że można ocenić rozwój jej efektywności poprzez obserwację zwiększania się odsetka reperacji ukrytych wraz z wiekiem. Zgodnie z tą hipotezą, wyniki zadania pokazały, że kolejne grupy wiekowe różnią się procentem odpowiedzi z reperacjami ukrytymi w strategii odraczania. O ile w przypadku dzieci średnio połowa zdań poprawiana była w ten sposób, o tyle w przypadku młodzieży było to prawie 60%, a w grupie dorosłych ponad 70%. Udało się więc potwierdzić, że rozwój kontroli procesu mówienia obejmuje wzrost sprawności wewnętrznej pętli monitorowania, umożliwiając większą kontrolę wypowiedzi jeszcze przed jej wyartykułowaniem.

Mając na względzie wpływ tempa mówienia na wyniki szybkości dokonywania interupcji w warunku błédnego podmiotu, postanowiono dodatkowo sprawdzić, na ile tempo mówienia mogło wpłynąć na liczbę reperacji ukrytych w warunku błédnego dopełnienia. Analiza korelacji cząstkowej między tempem mówienia (w warunku neutralnym) i odsetkiem reperacji ukrytych (w warunku błédnego dopełnienia) przy kontroli zmiennej „wiek” wykazała ujemny, istotny związek. Oznacza to, że im szybsze było tempo mówienia badanej osoby, tym rzadziej stosowała reperacje ukryte (przez odroczenie). Z tego względu stosunkowo wysoki odsetek reperacji ukrytych dokonanych przez dzieci należy przynajmniej częściowo przypisać temu, że zastosowana procedura faworyzowała osoby mówiące w niższym tempie. Jednocześnie analiza różnic w tempie mówienia między warunkiem neutralnym, a warunkiem błédnego dopełnienia nie wykazała, aby któraś z grup wiekowych bardziej spowalniała swoją artykulację dla wprowadzenia reperacji ukrytej. Strategia odraczania we wszystkich badanych grupach wiekowych wiązała się ze spowolnieniem tempa artykulacji o około 0,6 sylaby na sekundę. Biorąc jednak pod uwagę, że dzieci mówiły w najwolniejszym tempie, spadek ten powodował procentowo największe opóźnienie w tej grupie wiekowej w porównaniu do proporcjonalnych spadków tempa mówienia w pozostałych grupach wiekowych. Różnice te były jednak niewielkie.

Wyniki uzyskane w warunkach błędnego dopełnienia ponownie pokazują (zob. wyniki zadania *Monitorowanie sylab*), że rozwój mechanizmów kontroli mówienia prowadzi do wzrostu efektywności wewnętrznej pętli monitorowania (por. Sasisekaran i Weber-Fox, 2012). Próbuując uogólnić ten efekt na produkcję mowy w typowych sytuacjach komunikacyjnych, poza kontekstem specyficznie eksperymentalnym, można spodziewać się, że użytkownicy języka wraz z wiekiem mogą popełniać mniej jawnych przejęzyczeń, gdyż korekty coraz częściej dokonywać mogą się w planie przedartykulacyjnym. Poza tym, można spodziewać się, że wraz z wiekiem będzie występować przerwań i odtworzeń (objawy ukrytych reperacji przez odtworzenie, por. Levelt, 1989; Postma i Kolk, 1993) gdyż ten sam efekt może być osiąganym dzięki strategii odraczania. Niestety uzyskane w niniejszym badaniu wyniki nie mają swoich odpowiedników w analizach spontanicznych wypowiedzi u osób w podobnym wieku (dominują badania nad mową dzieci przed okresem szkolnym lub analizy mowy osób dorosłych), stąd trudno wskazać jest analogie do innych badań. Tym bardziej ciekawe jest spostrzeżenie Evans (1985), że związek częstości reperacji i wieku może mieć charakter odwróconej funkcji U-kształtnej. Według niej, początkowo wraz ze wzrostem umiejętności monitorowania wzrasta też częstość autonapraw, ale dzięki dalszemu rozwojowi umiejętności planowania mówienia i organizacji myślenia, korekty błędów są mniej potrzebne, a ich występowanie staje się rzadsze.

Chcąc wyjaśnić mechanizm rozwojowy, stojący za progresem obserwowanym w poprzecznej perspektywie omawianych wyników, należy ponownie odwołać się do teorii podwójnej pętli percepcyjnej. Zgodnie z modelem Levelta (Hartsuiker i Kolk, 2001; Hartsuiker i in., 2005; Levelt, 1989; Levelt i in., 1999;) wewnętrzna pętla monitorowania do swojego efektywnego działania konsumuje więcej zasobów niż kanał zewnętrzny, dzięki czemu wykrywa więcej błędów (szczególnie leksykalnych). Można więc przypuszczać, że dojrzewanie układu nerwowego i postępujący wraz z nim rozwój kompetencji językowej (Kurcz, 2005) oraz wzrost poziomu funkcji wykonawczych (Brzezińska i Nowotnik, 2012; Diamond, 2013; Józefacka-Szram, 2014; Woźniak-Prus i Gambin, 2011) przekłada się na większą skuteczność nie tylko w wykrywaniu błędów (Hanley i in., 2016), ale też w prewencyjnym zapobieganiu ich popełnianiu. Trudno jest jednak wskazać, czy wraz ze wzrostem zasobów pamięciowych zwiększa się np. zakres bufora artykulacyjnego, czy raczej zakres ten pozostaje niezmienny, a zwiększa się tempo przetwarzania zawartych w nim informacji. Być może za uzyskanymi wynikami stoją oba kierunki zmian.

Warto w tym miejscu zwrócić uwagę na to, że obserwowany w warunkach neutralnym zadania *Powtarzanie zdań* wzrost tempa mówienia wraz z wiekiem, może odzwierciedlać

ogólny trend rozwojowy, polegający na wzrastaniu szybkości przebiegu procesów umysłowych. Dotyczy to zresztą także omówionych wcześniej wyników zadań: *Nazywanie obrazków* oraz *Monitorowanie sylab*. W tym kontekście istotnym okazuje się pytanie o podłoże rozwoju kontroli procesu mówienia. Czy obserwowany w wynikach niniejszego badania progres jest zjawiskiem specyficznym przynależnym procesowi mówienia, czy też wiąże się z bardziej ogólnymi zmianami w kontroli poznawczej, wynikającymi z dojrzewania układu nerwowego i zdobywania doświadczenia. Inaczej mówiąc, czy wyróżnione aspekty kontroli procesu mówienia wykazują związki z poziomem funkcji wykonawczych? Temu problemowi poświęcona była druga grupa pytań badawczych.

#### **7.1.4 Kontrola procesu mówienia a poziom funkcji wykonawczych**

Wyniki dostarczone przez zadania psycholingwistyczne *Nazywanie obrazków*, *Monitorowanie sylab* i *Powtarzanie zdań*, pokazały, że każdy z wybranych aspektów kontroli procesu mówienia podlega zmianom rozwojowym obserwowanym poprzecznie w okresie od późnego dzieciństwa do dorosłości. Czas ten obfituje również w progresywne zmiany w zakresie kontroli poznawczej, gdyż to właśnie w okresie szkolnym (późne dzieciństwo i adolescencja) sprawność funkcji wykonawczych stopniowo zbliża się do poziomu osób dorosłych (Best i Miller, 2010; Józefacka-Szram, 2014). Dzięki temu zwiększają się możliwości sterowania celowym działaniem, także związanym z realizacją celów społecznych.

Podłożem rozwoju funkcji wykonawczych w dzieciństwie i adolescencji jest zarówno dojrzewaniem struktur mózgowych (szczególnie obszarów czołowych) oraz doświadczenie zdobywane w interakcjach ze środowiskiem fizycznym i społeczno-kulturowym (Brzezińska i Nowotnik, 2012; Diamond, 2013). Szczególnie intensywne zmiany w zakresie kontroli wykonawczej zachodzą w okresie adolescencji, kiedy to, w związku z intensyfikacją dojrzewania biologicznego, reorganizacji podlegają systemy funkcjonalne mózgu, będące neuronalną bazą funkcji wykonawczych. Do najważniejszych procesów w rozwoju układu nerwowego, prowadzących do osiągnięcia dorosłego poziomu kontroli poznawczej, umożliwiającej elastyczne i dowolnie kontrolowane zachowanie ukierunkowane na cel należą mielinizacja oraz przycinanie synaps (Luna, 2009). Ten pierwszy proces, nieprzerwanie postępujący od wczesnego dzieciństwa (choć w różnym tempie przebiega w poszczególnych obszarach kory mózgowej), zapewnia ciągły wzrost szybkości przesyłu impulsów nerwowych, zwiększając efektywność przetwarzania informacji. Z kolei przycinanie synaps poza wczesnym dzieciństwem ma swoją szczególną intensyfikację właśnie w okresie

adolescencji. Dzięki temu procesowi tracone są nieużywane połączenia synaptyczne, co w konsekwencji zwiększa wydajność obliczeniową, ponieważ obniża rozproszenie aktywności mózgu na rzecz bardziej sprawnego i wyspecjalizowanego przetwarzania informacji w obwodach regionalnych, typowego dla osób dorosłych (Gage i Baars, 2018; Luna, 2009).

Jeśli w tym samym czasie obserwuje się także progresywne zmiany w kontroli procesu mówienia, to istotnym staje się pytanie o możliwe związki między wzrastającym poziomem funkcji wykonawczych, a coraz sprawniej działającą kontrolą produkcji mowy. Jak wykazano w rozdziale 3., istnieją teoretyczne oraz empiryczne przesłanki, aby efektywność monitorowania produkcji mowy wiązać z bardziej ogólnymi mechanizmami kontroli poznawczej. Tym samym na podstawie obserwowanego w okresie od dzieciństwa do dorosłości wzrostu poziomu funkcji wykonawczych, powinno być możliwe prognozowanie progresu w zakresie kontroli procesu mówienia. Problemowi temu poświęcono kolejne pytania badawcze, do których następnie sformułowano hipotezy o związkach między poziomem funkcji wykonawczych a kontrolą selekcji leksykalnej, automonitorowaniem wewnętrznym oraz dokonywaniem reperacji. Hipotezy te weryfikowane były w modelu korelacyjnym na podstawie omówionych już wyników zadań psycholingwistycznych oraz wyników zadań mierzących funkcje wykonawcze: *Zadanie z flankerami*, *Zadanie n-wstecz*, *WCST (BCST)*. Przed przystąpieniem do testowania hipotez sprawdzono, czy zgodnie ze wstępnymi założeniami, funkcje wykonawcze badane wymienionymi testami będą wykazywać tendencje wzrostowe wraz z wiekiem. Założenia te potwierdziły się, a wyniki użytych zadań wykazywały systematyczny wzrost poziomu kontroli hamowania, odświeżania pamięci roboczej oraz giętkości poznawczej, różnicując tym samym badane grupy wiekowe<sup>183</sup>.

W pierwszej kolejności postanowiono zbadać związki między kontrolą selekcji leksykalnej a poziomem funkcji wykonawczych. Według Roelofsa (2005b; 2020a) mechanizmy weryfikujące wybrane do realizacji jednostki leksykalne wykazują związki z kontrolą wykonawczą, angażując w dużej mierze te same obszary korowe i podkorowe mózgu. W związku z tym postawiono sprawdzić, czy odporność selekcji leksykalnej na dystrakcję semantyczną zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w różnym wieku (hipoteza 2.1.1.) oraz czy wielkość efektu interferencji semantycznej maleje wraz ze wzrostem poziomu funkcji wykonawczych u osób w różnym wieku (hipoteza

---

183 O ile badane grupy wiekowe różniły się istotnie pod względem poziomu odświeżania pamięci roboczej (wyniki zadania 1-wstecz) o tyle poziom kontroli hamowania (wielkość kosztu konfliktu w zadaniu z flankerami) różnił się u dzieci i młodzieży oraz u dzieci i dorosłych (brak istotnych statystycznie różnic między młodzieżą i dorosłymi), natomiast poziom elastyczności poznawczej (wyniki zadania WCST) różnicował grupy dzieci i młodzieży z grupą dorosłych (brak istotnych statystycznie różnic między dziećmi i młodzieżą).

2.1.2.). Przyjętą miarą odporności selekcji leksykalnej na dystrakcję była różnica poprawności odpowiedzi między warunkiem neutralnym (bez dystrakcji) zadania *Nazywanie obrazków*, a warunkiem interferencji (z dystrakcją). Wziąwszy pod uwagę, że różnice między warunkami w grupie dorosłych i młodzieży okazały się nieistotne statystycznie pierwszą ze wspomnianych hipotez postanowiono sprawdzić jedynie w grupie dzieci. Analiza korelacji monotonicznej Spearmana nie wykazała istotnych związków między odpornością selekcji leksykalnej na dystrakcję u dzieci, a poziomem funkcji wykonawczych. Hipoteza 2.1.1. nie została więc potwierdzona w odniesieniu do żadnej składowej funkcji wykonawczych.

Analogicznie, tyle że na bazie wyników osób ze wszystkich badanych grup wiekowych, postanowiono zweryfikować drugą z powyższych hipotez. Analiza korelacji Spearmana przy kontroli wieku osób badanych (wiek traktowany jako kowariant) ponownie nie wykazała istotnych związków między wielkością efektu interferencji semantycznej a poziomem funkcji wykonawczych. Hipoteza 2.1.2. nie została więc potwierdzona w odniesieniu do żadnej składowej funkcji wykonawczych.

Brak wykazanych związków w obu analizach korelacji może świadczyć o tym, że selekcja leksykalna jest procesem niezależnym od centralnej kontroli poznawczej i przebiega w sposób typowy dla procesów zachodzących w zamkniętych modułach (zob. rozdział 3.). Tym samym wybór jednostek leksykalnych i związana z tym regulacja aktywacji przechodzącej między węzłami sieci semantycznej odbywałaby się w sposób niedostępny innym, zewnętrznym procesom kontrolnym (por. Meyer i in., 2007). Nie jest to wykluczone, zwłaszcza że pierwotny model Levelta (1989) zakładał do pewnego stopnia autonomiczność, hermetyczność informacyjną i wewnętrzny automatyzm działania formulatora (zob. rozdział 2.), odpowiedzialnego za selekcję jednostek leksykalnych. Możliwe zatem, że o ile monitorowanie produkcji mowy w przypadku wykrycia pomyłki korzysta z uniwersalnych funkcji kontrolnych (zlokalizowanych w korze przedniego zakrętu obręczy, Roelofs, 2020a), o tyle wczesne mechanizmy weryfikacyjne zaangażowane w selekcję leksykalną (i regulację poziomu aktywacji węzłów w sieci leksykalnej) są wyspecjalizowaną i wewnętrzną funkcją systemu produkcji mowy.

Istnieje jednak alternatywne wyjaśnienie, uwzględniające mimo wszystko ograniczony udział ogólnej kontroli wykonawczej w selekcji leksykalnej. W pierwszej kolejności należałoby spojrzeć dokładniej na wyniki cytowanego badania Shao i in. (2012). Otóż przedstawione przez badaczy związki między wynikami nazywania obrazków, a testami funkcji wykonawczych dotyczyły szybkości nazywania w warunkach bez dystrakcji. Nie

korelowano zatem wielkości efektu interferencji semantycznej z poziomem funkcji wykonawczych, lecz bezwzględny czas nazywania obrazków. Próba replikacji tych efektów na wynikach uzyskanych w badaniu własnym dała pozytywne rezultaty. Podobnie jak w badaniu Shao i in. (2012) udało się wykazać, że czas nazywania koreluje dodatnio z miarą hamowania (im mniejszy koszt hamowania, tym krótszy czas nazywania) oraz ujemnie ze wskaźnikiem odświeżania pamięci roboczej (im niższa sprawność funkcji pamięci roboczej, tym dłuższy czas nazywania). Można zatem przyjąć podobną do autorów badania interpretację, że zdolności odświeżania pamięci roboczej i kontroli hamowania wiążą się z szybkością nazywania obrazków.

Warto w tym miejscu zauważyć, że zgodnie z powyższym, selekcja leksykalna (oraz potencjalnie inne poziomy produkcji mowy zaangażowane w nazywanie obrazków) jeszcze przed wprowadzenia dodatkowej dystrakcji wykazuje związki z funkcjami wykonawczymi. Może to oznaczać, że już sam „szum” pochodzący z systemu leksykalnego generuje na tyle wysoki poziom dystrakcji, aby w proces produkcji mowy zaangażowana była kontrola wykonawcza. Zgadza się to z wynikami badania Abela i in. (2009), w którym obszary przedczołowe obu półkul wykazywały silną aktywację podczas zadania nazywania obrazków, zarówno kiedy osoba prawidłowo generowała wyrazy, jak i wtedy kiedy popełniała błąd (wówczas aktywność tych obszarów była jeszcze wyższa). Podobnie w badaniu Piai i in. (2014), podczas nazywania obrazków z dystrakcją swoją aktywność wykazywały neurony w lewym górnym zakręcie czołowym, którym przypisywana jest funkcja ogólnej kontroli uwagi (*domain-general attentional control*). Aktywność ta była największa w przypadku dystrakcji o charakterze semantycznym, ale występowała także w warunku z dystrakcją niesemantyczną (wyraz niepowiązany z nazwą obrazka) oraz w warunku bez dystrakcji (podane słowo towarzyszące obrazkowi było jego właściwą nazwą).

Na podstawie tych wyników można wywnioskować, że selekcja leksykalna odbywa się w warunkach nieustannej konkurencji (*competition*), która wymaga kontroli (weryfikacji) poziomu aktywacji węzłów w sieci leksykalnej. Zewnętrzna dystrakcja semantyczna, tylko nasila ten efekt ale nie powoduje zaangażowania innych obwodów funkcjonalnych, służących rozwiązaniu konfliktu między aktywowanymi jednostkami leksykalnymi. Tym samym związki funkcji wykonawczych z nazywaniem obrazków ujawniają się już w warunku prostego nazywania, które absorbuje większą część kontroli poznawczej, angażowanej także w zadaniach z manipulacją poziomem zewnętrznej dystrakcji.

Warto też w tym miejscu zwrócić uwagę na pewne ograniczenia metodologiczne. Ostatecznie efekt interferencji w nazywaniu wywoływany jest dystrakcją o charakterze

semantycznym, podczas gdy konflikt w zadaniu z flankerami nie jest rozstrzygany w oparciu o kryteria semantyczne, lecz percepcyjne (Eriksen i Eriksen, 1974). Oba zadania mogą więc angażować kontrolę hamowania, ale w specyficzny dla siebie sposób, przez co brak jest istotnych korelacji między kosztem konfliktu, a efektem interferencji semantycznej. Ponadto, jak sugerują Shao i in. (2013) wykazanie związków między funkcjami wykonawczymi (szczególnie kontrolą hamowania), a hamowaniem związanym z selekcją leksykalną w zadaniu nazywania z interferencją może być kwestią bardziej subtelnej analizy statystycznej. Według nich, odpowiednim materiałem do korelacji z funkcją hamowania selektywnego nie jest średni czas uzyskiwany przez osoby badane w całym zadaniu, ale dane z zestawienia czasów reakcji w poszczególnych segmentach (kwantylach) rozkładu. Ich zdaniem efekt interferencji powinien być obliczany dla każdego kwantyla osobno, a kwantyl z największymi różnicami (*the slope of the slowest segment*) powinien być dopiero brany pod uwagę jako miara hamowania selektywnego. Niestety ze względu na sposób kodowania odpowiedzi w niniejszym badaniu nie było możliwe przeprowadzenie analizy wykresu delta (różnic w poszczególnych kwantylach) i obliczenie efektu interferencji w osobnych kwantylach (segmentach) rozkładu według wzoru przedstawionego przez Shao i in. (2013). Tym samym nie jest możliwe sprawdzenie na zebranych danych ich założenia o związku hamowania selektywnego z najwolniejszym segmentem wykresu delty.

W kolejnym kroku podjęto weryfikację hipotez o związkach automonitorowania wewnętrznego z funkcjami wykonawczymi. Jak zauważyli Wheeldon i Levelt (1995) dodatkowe obciążenie zasobów poznawczych osób badanych zadaniem głośnego odliczania spowodowało niewielkie spadki poprawności oraz szybkości monitorowania fonemów. Efekt ten – choć niewielki – implikuje pytanie o to, w jakim stopniu efektywność automonitorowania procesu mówienia wiąże się poziomem kontroli wykonawczej, szczególnie pamięci roboczej. Aby odpowiedzieć na to pytanie postanowiono sprawdzić dwie hipotezy szczegółowe mówiące o tym, że dokładność monitorowania wewnętrznego zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych (hipoteza 2.2.1.) oraz, że szybkość monitorowania wewnętrznego zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych (hipoteza 2.2.2.). W celu weryfikacji hipotez przeprowadzono analizę korelacji cząstkowych testem Spearmana, między wskaźnikami sprawności automonitorowania procesu mówienia (wyniki zadania *Monitorowanie sylab*) a miarami funkcji wykonawczych, przy kontroli zmiennej „wiek”. Uzyskano wyniki częściowo potwierdzające założenie o związkach między efektywnością automonitorowania mowy, a poziomem funkcji wykonawczych. Zarówno wyniki kontroli hamowania, jak i odświeżania

pamięci korelowały z dokładnością monitorowania wewnętrznego (poprawność odpowiedzi).

Oznacza to, że

wyższy poziom tych składowych funkcji wykonawczych znajdował odzwierciedlenie w lepszym (bardziej poprawnym) wykonaniu zadania *Monitorowanie sylab*. Z kolei czas reakcji w tym zadaniu (szybkość monitorowania) korelował dodatnio z odsetkiem błędów perseweracyjnych w sortowaniu kart (WCST), co oznacza, że przy wyższym poziomie giętkości poznawczej wzrastała też szybkość monitorowania wewnętrznego. Na podstawie wartości współczynników korelacji ukazane związki można uznać za stosunkowo słabe.

Trudno w tym miejscu wskazać na ile uzyskane wyniki odzwierciedlają rzeczywisty udział tych funkcji wykonawczych w samym procesie monitorowania mowy, a na ile są specyficzne dla metodologii zadania *Monitorowanie sylab*, w którym odpowiedzi zbierane były poprzez wciśnięcie odpowiedniego przycisku na klawiaturze komputera. Niewykluczone, że np. zaangażowanie funkcji hamowania w większym stopniu wiązało się z kontrolą właściwego wyboru i decyzją, której ręki użyć, niż dokonywaniem inspekcji wewnętrznego planu mowy. Kolejne badania powinny uwzględnić ten problem, aby dać rozstrzygające rezultaty. Pewne światło na to zagadnienie może rzucić też analiza wyników weryfikujących kolejne hipotezy (2.3.1. oraz 2.3.2.) poświęcone związkom między reperacjami błędów mowy, a kontrolą wykonawczą.

Jak wspomniano, ze względu na bezpośrednią inspirację wynikami badania Wheeldon i Levelta (1995) najistotniejsze w testowaniu hipotez o związkach monitorowania wewnętrznego z funkcjami wykonawczymi było sprawdzenie na ile sprawność monitorowania wewnętrznego może współzależać od działania pamięci roboczej. Analiza korelacji cząstkowej między miarą aktualizacji pamięci (zadanie 1-wstecz), a dokładnością i szybkością monitorowania nie dała jednoznacznych rezultatów. Okazało się, że przy kontroli zmiennej „wiek” odświeżanie pamięci roboczej wykazuje związki jedynie z dokładnością monitorowania (poprawność wykrywania sylab), ale już nie z czasem potrzebnym na to efektywne monitorowanie (szybkość wykrywania sylab). Inaczej mówiąc, wraz ze wzrostem poziomu tej funkcji wykonawczej, wzrasta także dokładność monitorowania, ale nie jego szybkość.

W związku z tym, że nie udało się wykazać zakładanego związku między poprawnością zadania 1-wstecz a szybkością monitorowania wewnętrznego, postanowiono przeprowadzić dodatkową analizę korelacji, wykorzystując warunek zadania 2-wstecz, który bardziej różnicował wyniki w badanej próbie, ale okazał się za trudny dla większości dzieci i dużej części młodzieży. Z tego względu oraz biorąc pod uwagę fakt, że w zadaniu Wheeldon i

Levelta (1995) brały udział jedynie osoby dorosłe, do analizy wzięto jedynie dane osób z grupy dorosłych (usunięto odpowiedzi dwóch osób, które rozpoznano jako obserwacje nietypowe, zakłócające rozkład wyników). Tym razem uzyskano rezultaty spójne z postawioną hipotezą: wraz ze wzrostem poziomu funkcji odświeżania (poprawność odpowiedzi w zadaniu 2-wstecz), spadał czas potrzebny na efektywne monitorowanie wewnętrzne (wzrastała szybkość monitorowania).

Podsumowując, otrzymane wyniki potwierdzają, że monitorowanie wewnętrzne procesu mówienia wykazuje pozytywne związki z poziomem funkcji wykonawczych. Wzrost efektywności kontroli poznawczej w zakresie funkcji hamowania i odświeżania pamięci roboczej idzie w parze ze wzrostem dokładności monitorowania wewnętrznego. Ponadto w przypadku badanych osób dorosłych wyższy poziom funkcji odświeżania pamięci roboczej istotnie korelował ze skróceniem czasu, w jakim to monitorowanie przebiega. Szybkość monitorowania korelowała także istotnie z giętkością poznawczą w całej badanej próbie. Wyniki te wspierają więc założenie, że wysiłek (*effort*) wkładany w działanie wewnętrznej pętli monitorowania (Hartsuiker i in., 2005) może znajdować oparcie w bardziej ogólnych mechanizmach kontroli poznawczej. Choć w niniejszym badaniu przedmiotem monitorowania były sylaby, to według Nooteboom i Quené (2019) analogicznie działa monitorowanie przejęzyczeń. Przypomnijmy, że w modelu Levelta i in. (1999) „monitor”, pełni rolę „krytyka”, porównując reprezentacje użyte w produkcji mowy z reprezentacjami aktywowanymi przez system odbioru mowy. W przypadku wykrycia niezgodności (realna lub potencjalna pomyłka) sygnał błędu wysyłany jest do „kontrolera”, który modyfikuje realizację planu przez system produkcji mowy (Roelofs, 2020). Jednocześnie wykazane związki są słabe i niewystarczające, aby funkcjom wykonawczym przypisać nadrzędną rolę w zarządzaniu pracą „monitora” mowy.

Na podstawie uzyskanych danych – w związku z zastosowaną metodologią – nie można także orzec na ile wykazane związki między wykrywaniem sylab a funkcjami wykonawczymi, dotyczą wyłącznie monitorowania (aspekt kontroli produkcji mowy), a na ile innych procesów i czynności zaangażowanych w wykonanie zadania *Monitorowanie sylab* (zadania, w którym odpowiedzi zbierane były manualnie, a nie werbalnie). Uwzględnienie tego problemu jest zadaniem stojącym przed kolejnymi badaniami. Częściowo jednak może je rozwikłać analiza wyników zadania *Powtarzanie zdań*, w którym efekty monitorowania mowy ujawniały się w postaci przerywania, opóźniania i wznawiania artykulacji (bez konieczności udzielania odpowiedzi manualnych na klawiaturze). W związku z tym, że

warunki zadania pozwalały na pomiar szybkości dokonywania interupcji oraz częstość stosowania reperacji przez odraczenie (reperacje ukryte), postawiono hipotezy:

– Hipoteza 2.3.1. Szybkość dokonywania interupcji błędnych wypowiedzi zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w różnym wieku.

– Hipoteza 2.3.2. Częstość stosowania reperacji wewnętrznych przez strategię odraczenia zwiększa się wraz ze wzrostem poziomu funkcji wykonawczych u osób w różnym wieku.

Aby zweryfikować postawione hipotezy przeprowadzono analizę korelacji cząstkowych między szybkością dokonywania interupcji i częstością stosowania reperacji ukrytych a miarami funkcji wykonawczych, przy kontroli zmiennych „wiek” oraz „tempo mówienia<sup>184</sup>”. Okazało się, że spośród 3-czynnikowego modelu funkcji wykonawczych, jedynie funkcja kontroli hamowania wykazała istotne związki ze sposobem dokonywania reperacji.

Po pierwsze, szybkość dokonywania interupcji (mierzona liczbą sylab składających się na wypowiedź z reparandum) nie korelowała istotnie z żadną spośród zastosowanych miar funkcji wykonawczych. Nie można więc stwierdzić, że natychmiastowe lub odroczone przerywanie mówienia zależy w jakimkolwiek stopniu od funkcji wykonawczych (a przynajmniej nie wykazano takich zależności w tej części wariancji, która nie została wyjaśniona przez zmienne „wiek” i „tempo mówienia”).

Po drugie, częstość dokonywania reperacji ukrytych (strategia odroczenia) przy kontroli wieku oraz tempa mówienia uczestników wykazała istotne negatywne związki z miarą kontroli hamowania (pomiar kosztu konfliktu w zadaniu z flankerami). Oznacza to, że im wyższy poziom kontroli hamowania prezentowały badane osoby (coraz niższa wartość kosztu konfliktu), tym więcej reperacji ukrytych (przez odroczenie) były w stanie zastosować. Ponownie, uzyskany efekt był raczej niewielki, ale zgodny z przypuszczeniem, że reperacje ukryte inicjowane przez wewnętrzną pętlę monitorowania wymagają zasobów wykorzystywanych także przez kontrolę wykonawczą. Otrzymane wyniki pokazują, że wyższy poziom hamowania nie musi przejawiać się przerywaniem artykulacji (brak związku z szybkością interupcji), ale jej bardziej płynną formą kontroli w postaci spowolnienia lub zawieszenia, pozwalając tym samym na odroczenie egzekucji danej części wypowiedzi (np. błędnego wyrazu) na rzecz wprowadzenia korekty wewnętrznej. Interpretacja ta koresponduje z wynikami uzyskanymi przez Engelhardta i in. (2013) w badaniu, w którym niższy poziom kontroli hamowania związany był z występowaniem większej liczby niepłynności w postaci reperacji i rewizji szyku zdania. Być może, podobnie jak w niniejszym badaniu, osoby

---

184 W związku z wcześniejszym wykazaniem wpływu tempa mówienia na sposób dokonywania reperacji (zob. sekcja 6.4.2.1.), postanowiono włączyć tę zmienną – obok zmiennej „wiek” do modelu analizy korelacji, jako kowariant.

wykazujące się bardziej zaawansowaną kontrolą hamowania były w stanie dokonać niezbędnych reperacji jeszcze w buforze artykulacyjnym bez konieczności przerywania reparandum i jego restartu, a jedynie poprzez spowolnienie artykulacji. Niestety rejestrowane przez badaczy przejawy nie płynności obejmowały jedynie pauzy, powtórzenia i jawne reperacje, stąd trudno na ich podstawie wnioskować o częstotliwości reperacji ukrytych. Dalsze badania powinny uwzględnić wskaźniki niejawnych reperacji dokonywanych przez odroczenie, które zaproponowali Postma i Kolk (1993).

## 7.2 Ograniczenia przeprowadzonego badania

Przeprowadzone badanie ma szereg ograniczeń, które należy uwzględnić interpretując przedstawione wyniki i próbując dokonać generalizacji wyciągniętych wniosków.

Po pierwsze dobór osób do badania miał charakter celowo-losowy. Głównym kryterium doboru był wiek (oraz związany z nim etap edukacyjny), pozwalający przypisać uczestników do trzech grup wiekowych związanych z trzema różnymi okresami rozwojowymi: późne dzieciństwo, adolescencja, wczesna dorosłość. Kontrolowana była również zmienna płci osób badanych, aby zrównoważyć reprezentację kobiet i mężczyzn we wszystkich grupach wiekowych. Wszystkie osoby badane mieszkały (na stałe lub tymczasowo) na terenie Poznania, a dzieci i młodzież rekrutowano tylko z dwóch placówek edukacyjnych. Ogranicza to możliwość uznania próby za reprezentatywną, zwłaszcza, że nie były zbierane i kontrolowane dane na temat innych zmiennych (np. poziom inteligencji, status socjoekonomiczny itd.) niż te, które bezpośrednio wiązały się z postawionymi hipotezami badawczymi. W związku z powyższym należy ostrożnie uogólniać wnioski wyciągnięte na podstawie uzyskanych wyników.

Badanie, chociaż składało się z zadań o charakterze laboratoryjnym miało charakter quasi-eksperymentalny. Oznacza to, że zmienne niezależne w większości nie podlegały manipulacji, lecz pomiarowi. Z tego względu kontrola wariancji zmiennych niezależnych była znacznie ograniczona. W związku z tym, że badanie dotyczyło rozwojowych aspektów kontroli procesu mówienia, do głównych zmiennych niezależnych należały zmienna ilorazowa „wiek” oraz zmienna porządkowa „grupa wiekowa”. Mają one charakter zmiennych metrykalnych, które nie podlegają manipulacji eksperymentalnej. Możliwe jest jednak kontrolowanie tych zmiennych w planie podłużnym lub poprzecznym. W niniejszym badaniu wybrano plan poprzeczny, który ma swoje praktyczne zalety, ale jednocześnie ogranicza możliwości wnioskowania o procesach rozwojowych, sprowadzając je do porównania różnic indywidualnych związanych z wiekiem. Tym samym obserwowane różnice w poziomie kontroli procesu mówienia występujące między osobami badanymi w

trzech różnych grupach wiekowych informują jedynie o domniemanych zmianach rozwojowych, gdyż proces rozwoju u żadnej z badanych osób nie był obserwowany w czasie (każda z badanych osób badana była jednorazowo). To ograniczenie również należy uwzględnić przy interpretacji wyników i próbie ich generalizacji.

Ponadto poprzeczny schemat badania zakładał przedstawienie osobom badanym ze wszystkich grup wiekowych tych samych zadań. To wiązało się z koniecznością opracowania procedury na tyle przystępnej, aby mogły w niej uczestniczyć osoby najmłodsze, ale jednocześnie na tyle wymagającej, aby uzyskane wyniki były różnicujące ze względu na wiek osób badanych we wszystkich grupach. Ograniczenia z tego wynikające, to po pierwsze zawężenie porównań do okresu między późnym dzieciństwem a dorosłością, z wyłączeniem młodszych grup wiekowych. Ponadto, niektóre wyniki zdradzały wcześniejsze niż w dorosłości osiągnięcie efektu „sufitu”, przez co nie udało się w tych wypadkach wystarczająco wykazać zróżnicowania między grupą adolescentów i dorosłych (np. porównanie odporności selekcji leksykalnej na dystrakcję, czy kontroli hamowania). Poza tym, w zadaniu *n-wstecz* odpowiadanie w warunku *2-wstecz* okazało się za trudne dla badanych dzieci i większości adolescentów, dlatego do pomiaru odświeżania pamięci roboczej wykorzystano wyniki zadania *1-wstecz*, które z kolei w niewielkim stopniu różnicowało odpowiedzi w obrębie grupy osób dorosłych.

Podobnie do zmiennej „wiek”, zmienne niezależne związane z kontrolą poznawczą podlegały pomiarowi (a nie eksperymentalnej manipulacji) przez co potencjalne związki funkcji wykonawczych z kontrolą procesu mówienia sprawdzane były w schemacie korelacyjnym. Tym samym kierunkowa interpretacja o udziale funkcji wykonawczych w kontroli procesu mówienia (uznanie, że monitorowanie mowy i inne mechanizmy kontrolne polegają na funkcjach wykonawczych) jest nieprecyzyjna. Korelacja wskazuje na występowanie związku między zmiennymi, ale nie przesądza o kierunku wpływu przyczynowo-skutkowego. Dopóki więc badanie nie manipuluje poziomem zmiennej niezależnej (manipulowanie dostępnymi zasobami kontroli poznawczej, np. w paradygmacie zadania podwójnego) interpretacje przyczynowe są mocno ograniczone. Z tego względu wyniki niniejszego badania odnoszące się do związków kontroli procesu mówienia z funkcjami wykonawczymi sugerują co najwyżej, że procesy te współdzielą pewną pulę zasobów poznawczych lub są ze sobą powiązane w jakiś jeszcze inny sposób. Dalej idące wnioski mają charakter spekulacji wynikających z przesłanek teoretycznych oraz empirycznych zawartych w omówionej literaturze. Uzyskane w niniejszym badaniu wyniki nie mogą być jednak w tej kwestii rozstrzygające.

Ponadto, wykazane związki między funkcjami wykonawczymi a wybranymi aspektami kontroli poznawczej okazały się stosunkowo słabe. Częściowo odpowiadać może za to zastosowanie nieparametrycznego testu korelacji Spearmana, który w sytuacji liniowego związku między zmiennymi może mieć mniejszą moc niż analiza korelacji Pearsona. Za niskie wartości współczynników korelacji może odpowiadać też silny związek analizowanych zmiennych z kowariantem w postaci „wieku”. Z jednej strony, kontrola wpływu wieku była konieczna, aby zbadać niezależne od niego związki między funkcjami wykonawczymi a kontrolą procesu mówienia. Z drugiej jednak strony, odebrało to niemalą część ich wspólnej wariancji, wyjaśnianą przez zmienną kowariantną. Niewykluczone więc, że przy zastosowaniu innych metod statystycznych na analogicznym materiale empirycznym, ujawnione związki osiągnęłyby wyższą wartość lub więcej związków osiągnęłoby przyjęty poziom istotności.

Pewne ograniczenia przeprowadzonego badania wynikają także z zastosowanych metod pomiarowych. Choć wybór zadań do wykonania przez osoby badane podyktowany był zamiarem przeprowadzenia precyzyjnej mikroanalizy poszczególnych aspektów kontroli procesu mówienia, to jednocześnie krytycznie należy spojrzeć na trafność zewnętrzną zastosowanych metod. Ogólna uwaga dotyczy wszystkich zastosowanych zadań psycholingwistycznych (*Nazywanie obrazków*, *Monitorowanie sylab*, *Powtarzanie zdań*), które wymagały od osób badanych produkcji mowy (na głos lub bezgłośnie), ale poza kontekstem rzeczywistej komunikacji z drugim człowiekiem. Analizowane więc były czynności lokucyjne z pominięciem ważnego dla mowy aspektu czynności illokucyjnych i perlokucyjnych (por. Austin, 1962/1993; Clark, 1996). Zawsze w przypadku tego typu metodologii należy rozważyć na ile dane zadanie odzwierciedla czynności podejmowane przez typowych użytkowników języka w codziennych rozmowach. Ponadto, każde z zastosowanych zadań ma swoje własne ograniczenia, które wymagają komentarza.

1. *Nazywania obrazków* – jest metodą, która angażuje wszystkie piętra produkcji mowy, przy jednocześnie wysokim poziomie kontroli treści, która jest realizowana (od konceptualizacji po artykulację). Dzięki temu możliwy jest precyzyjny pomiar odpowiedzi pod względem ich poprawności oraz czasu reakcji. Dodatkowo możliwe jest aplikowanie różnych bodźców, które mogą wpłynąć na przebieg procesu generowania mowy (nazwy obrazka). Zewnętrzna dystrakcja semantyczna zastosowana w niniejszym badaniu (SOA = -100) utrudnia proces selekcji leksykalnej, stąd spadek poprawności nazywania oraz opóźnienie w czasie reakcji występujące pod wpływem bodźców dystrakcyjnych interpretowane było w

kategoriach odporności selekcji leksykalnej na dystrakcję. Podobnie w kategoriach skuteczności selekcji leksykalnej interpretowano wyniki nazywania bez dystrakcji, zwłaszcza że występowały różnice międzygrupowe w tym zakresie wynikające z wieku. Należy jednak zauważyć, że zarówno spadek poprawności, jak i wydłużenie czasu nazywania nie muszą wynikać jedynie z przebiegu selekcji leksykalnej, ale mogą mieć swoje źródło na innych poziomach produkcji mowy. Z tego względu – jak to zaznaczono wcześniej – nie można różnic w czasie nazywania utożsamiać wprost z różnicami w selekcji leksykalnej. Przyjęta zatem interpretacja wyników zadania nazywanie obrazków, głównie w odniesieniu do selekcji leksykalnej ma duże ograniczenia. Inne ograniczenie wynika z tego, że dystrakcja semantyczna podawana jest w sposób systematyczny. Choć zadbano o to, aby próby pochodzące z obu warunków (neutralnego i interferencji) następowały po sobie w sposób losowy, to jednak osoby badane mogły spodziewać się, a tym samym nastawić obronnie przed dystrakcją. Prawdopodobnie z tego względu interferencja między obrazkiem i wyrazem dystrakcyjnym objawia się głównie w postaci spowolnienie nazywania, nie powodując nasilenia przejęzyczeń, które z kolei w naturalnych sytuacjach mówienia, często mają swoje źródło w zewnętrznej dystrakcji (Harley, 1984; 2001). Inne ograniczenia związane z metodą nazywania obrazków, omówiono już wcześniej.

2. *Monitorowanie sylab* – polegało na wskazaniu przy pomocy odpowiedniego przycisku na klawiaturze komputera, czy podana sylaba docelowa znajduje się w strukturze nazwy obrazka wyświetlanego na ekranie. Dzięki temu możliwy był precyzyjny pomiar poprawności wykonania oraz czasu reakcji, przy podobnym jak w przypadku *Nazywania obrazków* poziomie kontroli treści produkcji mowy. Jednocześnie zadanie to w niewielkim stopniu przypominało pracę „monitora” w naturalnej sytuacji konwersacyjnej, w której wypowiedź sprawdzana jest pod wieloma aspektami (zgodność z intencją, poprawność gramatyczna, stylistyczna, dostosowanie przekazu do możliwości odbioru rozmówcy itp.) Choć Nootboom i Quené (2019) wykazali, że monitorowanie wewnętrzne w celu wykrywania błędów oraz monitorowanie wewnętrzne w celu wykrywania fonemów są procesami porównywalnymi, to jednak brali oni pod uwagę jedynie monitorowanie zaangażowane w wykrywanie błędów indukowanych eksperymentalnie przez łamańce językowe (*tongue twisters*). Tym samym pytanie o możliwość generalizacji wyników z zadania *Monitorowanie sylab* na monitorowanie procesu mówienia w warunkach realnej konwersacji pozostaje otwarte. Pytanie to jest zasadne jeszcze pod względem tego, jaki poziom świadomości

metajęzykowej wymagany jest w monitorowaniu sylab. Niewykluczone, że poza rzeczywistym zaangażowaniem procesu monitorowania mowy (tożsamego z monitorowaniem spontanicznych wypowiedzi), wymaganie od osób badanych, aby sekwencjonowały generowane wyrazy, dzieląc je na sylaby, wykorzystuje taki poziom wiedzy językowej (świadomości metajęzykowej), który uzależniony jest od wtórnych zjawisk nauki czytania i pisania (por. Krasowicz-Kupis, 2004; Kwarciak, 1995). Poza tym brak kontroli materiału językowego pod względem parametrów statystycznych (np. frekwencja użytych w badaniu wyrazów i sylab) może rzutować na uzyskane wyniki, co zostało szczegółowo omówione wcześniej w dyskusji wyników.

3. *Powtarzanie zdań* – jest zadaniem nawiązującym do paradygmatu wywoływania reperacji (*repair elicitation paradigm*). Należy jednak zauważyć, że korekty dokonywane przez osoby badane nie dotyczą rzeczywistych błędów w planowaniu lub programowaniu mowy, ale niezgodności między treścią wypowiedzi (opis obrazka) a treścią opisywanego obrazka, przy czym niezgodność ta jest wynikiem manipulacji eksperymentalnej (van Wijk i Kempen, 1987; por. Boland i in., 2005). Dzięki temu obserwować można zachowania werbalne typowe dla korekt spontanicznych błędów mowy (por. Levelt, 1983; Postma i Kolk, 1993): przerwanie reparandum i wprowadzenie reperacji przez restart wypowiedzi lub spowolnienie artykulacji w celu wprowadzenia reperacji ukrytej (strategia odraczania). Ponownie jednak sytuacja eksperymentalna odbiega od naturalnych warunków mówienia, w których pomyłki pojawiają się spontanicznie i incydentalnie, przez co ich korekty także nie są działaniem zamierzonym (mowa nie jest rozpoczynana w celu jej poprawienia). Z tego względu, pomimo zalet wynikających z kontroli warunków zadania, możliwość uogólnienia uzyskanych przy jego pomocy wyników jest ograniczona i wymaga ostrożności.

Podobne uwagi można poczynić względem trafności zewnętrznej zadań mierzących poziom funkcji wykonawczych. Istotnym ograniczeniem niniejszego badania jest wykorzystanie pojedynczych zadań do oceny poziomu poszczególnych funkcji wykonawczych. W literaturze przedmiotu podnosi się problem „zanieczyszczenia” (kontaminacji) wyników danego zadania różnym udziałem w jego wykonaniu poszczególnych funkcji (Diamond, 2013). Z tego względu preferowanym rozwiązaniem jest wykorzystanie do pomiaru każdej odrębnej funkcji przynajmniej dwóch lub więcej zadań, aby w analizie statystycznej opartej na modelowaniu równań strukturalnych uwzględnić zmienne latentne, odpowiadające „czystym” funkcjom wykonawczym (por. Miyake i in. 2000; Lehto i in., 2003). Pociąga to za sobą szereg

komplikacji (np. wydłużenie czasu badania wynikające ze zwiększenia liczby zadań, znaczne zwiększenie próby itp.), na które nie zdecydowano się podczas projektowania badania. Z tego względu, interpretując uzyskane wyniki, należy mieć na uwadze to, że *Zadanie z flankerami*, *Zadanie n-wstecz* oraz *WCST (BCST)* w różnym stopniu angażują wszystkie funkcje kontroli wykonawczej, a pojedynczy wynik danego zadania nie jest uzależniony wyłącznie od poziomu jednej, izolowanej funkcji wykonawczej.

### 7.3 Implikacje dla dalszych badań i praktyki

Uzyskane w badaniu wyniki pozwalają uznać, że modele produkcji mowy, w tym – przyjęty w niniejszej pracy jako wyjściowy – model Levelta (1989; 1998; 1999a; 1999b; Levelt i in., 1999; Roelofs, 2005b; 2020a), które tworzone były w oparciu o dane z innych języków niż język polski, adekwatnie opisują sposób produkcji i kontroli mowy przez rodzimych użytkowników języka polskiego (por. Frydrychowicz, 1999). Ponadto wykazano, że uniwersalność tych modeli nie obejmuje tylko produkcji mowy przez osoby posługujące się różnymi językami, ale także przez osoby w różnym wieku. Tym samym potwierdzono wcześniejsze ustalenia, że w organizacji systemu produkcji mowy u dzieci i dorosłych występują raczej różnice ilościowe niż jakościowe (Budd i in., 2011; Jerger i in, 2002). Dotyczy to jednak dzieci w okresie późnego dzieciństwa, a więc mających za sobą okres najintensywniejszych zmian w zakresie nabywania języka (Kurcz, 2005). Tym samym badanie produkcji i kontroli mowy przez dzieci młodsze w okresie wczesnego i średniego dzieciństwa jest ważnym kierunkiem dalszych badań. Jest to o tyle istotne, że „dorosłe” modele produkcji mowy mogą nie być adekwatne dla wyjaśnienia tego jak zorganizowane jest mówienie małych dzieci (por. Levelt, 1998; Levy 1999; Rispoli, 2003).

Ponadto przeprowadzone badanie nie pozwala na rozstrzygnięcie kwestii spornych różniących dominujące obecnie modele produkcji mowy. Chociaż za punkt wyjścia przyjęto model Levelta, uzyskane wyniki dają się także zinterpretować w oparciu o model koneksjonistyczny Della (1986; Nozari i in, 2011; Nozari, 2020). Dalsze badania powinny zatem zaangażować na tyle precyzyjną metodologię, aby możliwa była falsyfikacja hipotez różnicujących konkurencyjne modele.

Poza tym, przeprowadzone badanie miało charakter laboratoryjny, przez co stawiane przed osobami badanymi zadania tylko w niewielkim stopniu przypominały naturalne sytuacje mówienia, związane z komunikacją interpersonalną. Niewątpliwie tak wysoki poziom kontroli warunków, w których odbywa się produkcja mowy uzupełnia dane uzyskane z naturalnych obserwacji (dzienniczki mowy, rejestracje spontanicznych przejęzyczeń i ich

korekt), ale ogranicza możliwość generalizacji wyników i zwiększa ryzyko uzyskania artefaktów. Potrzeba zatem więcej badań w kontrolowanych warunkach, ale angażujących proces mówienia w realnych sytuacjach komunikacyjnych.

Z drugiej strony, chcąc lepiej zrozumieć związki łączące kontrolę procesu mówienia z funkcjami wykonawczymi potrzebne są badania *stricte* eksperymentalne, w których poziom zaangażowanej kontroli wykonawczej podczas mówienia podlegałyby manipulacji. Tym samym konieczne jest opracowanie metodologii pozwalającej na bardziej precyzyjny poziom kontroli zmiennych, analizowanych także w niniejszym badaniu. Obiecujący jest tu także kierunek badań wykorzystujących metody neuroobrazowania.

Przeprowadzone badanie miało charakter poprzeczny, w czym podobne było do wielu cytowanych w niniejszej pracy badań. Uzupełnienie tej perspektywy o badania podłużne, szczególnie w grupie dzieci pozwoliłoby lepiej uchwycić dynamikę zmian rozwojowych oraz wykreślić trajektorie rozwoju poszczególnych aspektów produkcji mowy oraz jej monitorowania. Ponadto badania nad rozwojem kontroli procesu mówienia w kontekście zmian w obrębie funkcji wykonawczych powinny także objąć starsze grupy wiekowe. Jak pokazali np. Higby i in. (2019) efektywność funkcji wykonawczych u osób w wieku senioralnym pozwala przewidywać szybkość i poprawność nazywania obrazków. Tego typu analizy mogły by uzupełnić bogatą już literaturę na temat mówienia osób starzejących się.

Warto także zauważyć, że uzyskane odpowiedzi na postawione pytania badawcze nie wyczerpują podjętego tematu, lecz inspirują kolejne pytania i otwierają pola do dalszych dociekań naukowych. Kilka z nich zostanie wymienionych poniżej.

1. Jakie mechanizmy powodują, że przy silnym bodźcowaniu dystrakcyjnym kontekstem semantycznym wydłuża się czas nazywania oraz pogarsza poprawność (w przypadku dzieci), ale nie zwiększa się odsetek przejęzyczeń. Czy ma to swoje źródło w mechanizmach kontroli selekcji leksykalnej, czy w monitorowaniu efektów tej selekcji (zob. dyskusja między Dhooge i Hartsuiker, 2010; 2013 oraz Roelofs i in., 2011)? Jakie warunki są konieczne, aby dystrakcja semantyczna mogła indukować przejęzyczenia na poziomie leksykalnym (błędna selekcja leksykalna) znane z obserwacji spontanicznych przejęzyczeń?
2. Jakie poziomy produkcji mowy rzeczywiście trwają dłużej w nazywaniu obrazków przez dzieci w porównaniu ze starszymi grupami wiekowymi? Czy jest to głównie selekcja leksykalna, czy raczej proporcjonalnie dłużej trwa każdy z etapów produkcji mowy?

3. Jak różne parametry statystyczne języka (frekwencja wyrazów, frekwencja sylab, długość trwania fonemów, prawdopodobieństwo fonotaktyczne itp.) wpływają na szybkość monitorowania elementów subleksykalnych? Czy parametry te mają wpływ na skuteczność monitorowania i korygowania błędów mowy (por. Bręński, 2015)?
4. Od czego zależy preferencja płynności lub poprawności w dokonywaniu reperacji błędów mowy? Czy wskaźnikiem wyższego poziomu kontroli procesu mówienia jest szybkie korygowanie błędów mowy, czy raczej opóźnienie reperacji na rzecz zachowania płynności procesu mówienia?
5. Jakie obiektywne wskaźniki reperacji ukrytych przyjąć w analizie mowy spontanicznej?
6. Jaki jest udział funkcji wykonawczych, szczególnie funkcji hamowania w kontroli selekcji leksykalnej w różnym stopniu obciążonej czynnikami dystrakcyjnymi?
7. Czy wykazane związki funkcji wykonawczych z monitorowaniem sylab mają swoje przełożenie na monitorowanie błędów mowy oraz innych istotnych parametrów wypowiedzi podlegających kontroli „monitora”?
8. W jakim stopniu relatywnie niższy poziom funkcji wykonawczych ujawnia się w procesie mówienia? Czy zaburzenia płynności mówienia (np. jąkanie) można wyjaśnić zakłóceniami w procesie monitorowania mowy? Jeśli tak, to jakie czynniki leżą o podłoża tych zakłóceń? Czy potencjalnie niższy poziom funkcji wykonawczych można uznać za pierwotną przyczynę tych zakłóceń?

Ostatnie z wymienionych pytań mają dodatkową wartość, gdyż dotyczą problemów silnie związanych z praktyką diagnozy i terapii zaburzeń mowy. Choć przeprowadzone badanie miało charakter podstawowy, wyciągnięte na podstawie uzyskanych wyników wnioski, pozwalają także na sformułowanie implikacji praktycznych.

Po pierwsze, ujawnione w badaniu różnice między osobami w różnym wieku miały charakter ilościowy, pokazując że organizacja produkcji mowy i jej kontrola nie różnią się w swojej naturze u dzieci wczesnoszkolnych, młodzieży i młodych dorosłych. Jednocześnie udało się zaobserwować, że te ilościowe różnice w tempie przetwarzania materiału językowego mogą przekładać się na jakościowe efekty w płynności mowy. Dystrakcja semantyczna w zadaniu nazywania obrazków skutkowałą nie tylko typowym efektem interferencji polegającym na opóźnieniu nazywania, ale powodowała także wzrost pauz wypełnionych (u dzieci i młodzieży) oraz spadek poprawności nazywania (u dzieci). Pokazuje to, jak specyficzne warunki, np. presja czasowa, czy obecność dystraktorów może działać

destrukcyjnie na proces produkcji mowy, szczególnie u dzieci. Choć biorące udział w badaniu dzieci nie miały diagnozy zaburzeń mowy oraz były w wieku, w którym zazwyczaj nie obserwuje się już rozwojowej nie płynności mówienia (Tarkowski, 1992), specyficzne warunki eksperymentalne spowodowały u nich trudności w generowaniu słów (nazwy obrazków), które – jeśli wystąpiłyby w mowie spontanicznej – znacznie utrudniłyby sformułowanie płynnej wypowiedzi. To z kolei każe zastanowić się w jakich warunkach przebiega nabywanie języka i kształtowanie kompetencji językowej przez dzieci w młodszym wieku. Jakie działania są podejmowane przez rodziców, wychowawców, terapeutów, aby naturalnie pojawiające się zakłócenia płynności mowy, związane z rozwojem językowym (szczególnie w aspekcie nabywania systemu gramatycznego, który modyfikuje proces selekcji leksykalnej aspektami kodowania gramatycznego, por. Levelt, 1989; 1998) nie utrwały się w postaci jąkania? Wydaje się, że zapewnienie odpowiednich warunków dla niezakłóconego budowania wypowiedzi może być właściwą strategią wyjściową. Choć korekty ze strony otoczenia stanowią cenne źródło sprzężenia zwrotnego dla dziecięcych, niedoskonałych konstrukcji werbalnych, ciągle poprawianie, czy podpowiadanie poprawnych form językowych może wprowadzać dodatkowe zakłócenia, utrudniające jeszcze bardziej proces generowania wypowiedzi. Co więcej, jak wykazali Vasić i Wijnen (2005), zwracanie uwagi na symptomy nie płynności mówienia może je dodatkowo nasilać, przekierowując pracę „monitora” na ich restrykcyjne wykrywanie i korygowanie.

Choć analiza tempa mówienia, nie była przedmiotem zainteresowania badawczego w niniejszym projekcie, to wyniki zadania *Powtarzania zdań* ujawniły, że wolniejsze tempo mówienia (mierzone w warunkach neutralnych) sprzyja szybszemu przerywaniu błędnych wypowiedzi (krótsze reparandum w warunkach błędnego podmiotu) oraz zwiększa szansę na wprowadzenie reperacji ukrytej (w warunkach błędnego dopełnienia). Choć nie można wykluczyć, że do pewnego stopnia jest to efekt wynikający z zastosowanej metodologii (zob. sekcja 5.4.2.1.3.), to warto zastanowić się nad rolą tempa mówienia w regulowaniu płynności wypowiedzi. Wydaje się, że praca nad tym aspektem procesu mówienia, szczególnie trenowanie monitorowania tempa mówienia i jego świadomej kontroli może przynieść korzyści dla usprawniania płynności mowy, co ma już swoje zastosowanie w standardach terapii logopedycznej (Bijak i Kamińska, 2021; Woźniak, 2008).

Głównym aspektem przeprowadzonych analiz były związki kontroli procesu mówienia z funkcjami wykonawczymi w perspektywie rozwojowej. Choć pozytywna weryfikacja części hipotez o istnieniu takich związków, przeprowadzona była w planie korelacyjnym, można wyprowadzić pewne wnioski także dla działalności praktycznej. Po

pierwsze, diagnoza zaburzeń w zakresie produkcji mowy (jąkanie, gielkot, SLI, afazja itp.) nie powinna ograniczać się jedynie do testów typowo słownych, ale powinna także uwzględniać sprawność funkcji wykonawczych (por. Lambon i in., 2007). Podobnie, oddziaływanie terapeutyczne mogłoby uwzględnić szerszy kontekst funkcjonowania poznawczego, włączając pracę nad usprawnieniem funkcji wykonawczych. Możliwe, że to niespecyficzne działanie ukierunkowane na funkcje zarządzające planowaniem i realizacją celowego działania, mogłyby przynieść także korzyści dla usprawnienia procesu mówienia. Analogicznie, diagnoza deficytów w zakresie funkcji wykonawczych mogłaby pociągać za sobą praktykę objęcia oceną (a w razie konieczności wsparciem terapeutycznym) także funkcji językowych, w tym specyficznie związanych z kontrolą procesu mówienia. Oczywiście konieczne są dalsze badania podstawowe i stosowane, aby uprawomocnić zalecenia tego typu, gdyż wykazane w niniejszym projekcie związki są zbyt słabe i nie mogą stanowić wystarczającej podstawy dla konkretnych wskazań praktycznych.

## 7.4 Podsumowanie

Opisany projekt badawczy miał na celu analizę wybranych aspektów kontroli procesu mówienia pod względem ich zmienności wynikającej z wieku oraz potencjalnych związków z funkcjami wykonawczymi. Badanie miało charakter quasi-eksperymentalny i zostało przeprowadzone w planie poprzecznym, obejmując osoby badane w trzech grupach wiekowych: dzieci wczesnoszkolne, adolescentów i młodych dorosłych. Punktem wyjścia dla sformułowania hipotez badawczych były współczesne modele produkcji mowy (Harley, 2001; Levelt, 1999a; Vigliocco i Hartsuiker, 2005), a szczególnie model Levelta (1989; 1999b; Levelt i in., 1999) i związana z nim teoria podwójnej pętli percepcyjnej w monitorowaniu mowy (Hartsuiker, 2007; Hartsuiker i Kolk, 2001; Hartsuiker i in., 2005; Levelt, 1983; 1999b; Roelofs, 2005a; 2020a; Schiller, 2005).

Wyniki zastosowanych zadań psycholingwistycznych pozwalają stwierdzić, że w okresie od późnego dzieciństwa, przez dorastanie do wczesnej dorosłości kontrola procesu mówienia zwiększa swoją efektywność. Wyróżnione aspekty produkcji i kontroli mowy: skuteczność selekcji leksykalnej, odporność selekcji na dystrakcję, monitorowanie wewnętrzne, czy częstość stosowania reperacji ukrytych wykazują wraz z wiekiem zmiany progresywne. Wykazany w planie poprzecznym rozwój kontroli procesu mówienia ma charakter ilościowy, co pozwala stwierdzić, że produkcja mowy u dzieci w okresie późnego dzieciństwa jest zbliżona w swojej organizacji i przebiegu do produkcji mowy u adolescentów i osób dorosłych (por. Budd i in., 2011; Jerger i in., 2002). Jednocześnie wykazano, że te

ilościowe różnice mogą skutkować jakościowymi rezultatami, gdyż w różnym stopniu wpływają na płynność w wyartykułowaniu właściwego wyrazu (opóźnienie czasu reakcji w nazywaniu obrazków, poprzedzenie nazwy obrazka pauzą wypełnioną, brak podania nazwy obrazka).

Przeprowadzone badanie w wyniku zastosowanej manipulacji eksperymentalnej (manipulowanie kontekstem semantycznym, miejscem sylaby docelowej w strukturze monitorowanych wyrazów; miejscem błędu w strukturze powtarzanego zdania) potwierdziło, że efekty znane z badań nad użytkownikami innych języków, obecne są także w procesie produkcji mowy użytkowników języka polskiego. Zarówno efekt interferencji semantycznej, jak i przyrostowy charakter monitorowania elementów subleksykalnych (syłab) przyjęły kierunek i rozmiar porównywalny z dotychczasowymi doniesieniami. Co istotne efekty te szczególnie uwidoczniły się w najmłodszej grupie wiekowej, pozwalając na bardziej jednoznaczne interpretacje wyników, niż gdyby badanie dotyczyło wyłącznie osób dorosłych<sup>185</sup>.

Pewną innowacją w stosunku do dotychczasowych badań było zastosowanie zmodyfikowanego paradygmatu wywoływania reperacji (por. Boland i in., 2005; van Wijk i Kempen, 1987) do indukowania reperacji ukrytych (strategia odroczenia). Zgodnie z założeniami okazało się, że wraz z wiekiem zwiększa się potencjał zastosowania tego typu reperacji w radzeniu sobie z błędami wykrywanymi przez wewnętrzną pętlę monitorowania. Jest to o tyle istotne, że w badaniach spontanicznych przejrzyczeń i ich korekt, jedynym przejawem wykrywania ich jeszcze przez wewnętrzną pętlę monitorowania są tak zwane „preperacje” (*prepairs*) (Hartsuiker, Kolk i Martensen, 2005; por. Levelt, 1983; 1999; Postma i Kolk, 1993). Przy użyciu tego paradygmatu badawczego nie udało się natomiast jednoznacznie udowodnić, że szybkość dokonywania reperacji jawnych (indukowanych przez zewnętrzną pętlę monitorowania) zwiększa się wraz z wiekiem. Ujawniono natomiast, że istotną rolę w szybkości dokonywania reperacji (przynajmniej w ramach zastosowanej metodologii zadania badawczego) odgrywa tempo mówienia. Osoby mówiące w wolniejszym tempie przerywały reparandum po wypowiedzeniu mniejszej liczby sylab, a także były w stanie więcej błędów poprawić w strategii odraczania. Trudno uogólnić tę obserwację na warunki naturalnej konwersacji i monitorowania spontanicznej mowy, tym bardziej, że uzyskany efekt może mieć charakter artefaktu wynikającego z zastosowanej procedury. Tym

---

185 W przypadku osób dorosłych zewnętrzna dystrakcja semantyczna w nazywaniu obrazków nie powodowała spadku poprawności odpowiedzi ani nasilenia pauz wypełnionych. W tej grupie wiekowej różnice między szybkością monitorowania wyrazów 2 i 3 sylabowych były nieistotne statystycznie.

niemniej, obserwacja ta jest ważna z perspektywy dalszych badań psycholingwistycznych, gdyż pokazuje, że tempo mówienia powinno być istotną zmienną podlegającą kontroli badacza.

Szukając potencjalnych wyjaśnień dla obserwowanych zmian rozwojowych w kontroli procesu mówienia postanowiono sprawdzić, w jakim stopniu poziom tej kontroli może wynikać ze związków z funkcjami wykonawczymi. Przeprowadzone analizy korelacji z uwzględnieniem wieku osób badanych, jako kowariantu ujawniły istniejące relacje między wynikami zadań psycholingwistycznych, a wynikami zadań mierzących poziom funkcji hamowania, odświeżania pamięci i giętkości poznawczej. Po pierwsze wbrew oczekiwaniom odporność selekcji leksykalnej oraz wielkość efektu interferencji w nazywaniu obrazków nie wykazały związków z funkcjami wykonawczymi. Tym niemniej czas nazywania obrazków bez dodatkowej dystrakcji okazał się korelować z funkcją hamowania i odświeżania, co stanowi replikację wyników uzyskanych przez Shao i in. (2012). Pozwala to na wysunięcie przypuszczenia, że mechanizmy kontrolne zaangażowane w planowanie i programowanie mowy korzystają z ogólnej puli zasobów kontroli poznawczej nawet bez zewnętrznej dystrakcji (por. Abel i in., 2009; Piai i in., 2014). Również monitorowanie wewnętrzne (pomiar w zadaniu *Monitorowanie sylab*) wykazało pewne istotne związki z wszystkimi składowymi funkcjami wykonawczymi. Konieczne są jednak dalsze badania nad ustaleniem zakresu w jakim kontrola wykonawcza może mieć udział w monitorowaniu procesu mówienia. Jak wynika z analizy sposobów dokonywania reperacji, stosowanie korekt ukrytych (przez odroczenie), polegających na spowolnieniu lub chwilowemu zawieszeniu artykulacji istotnie koreluje z funkcją kontroli hamowania. Inaczej mówiąc, osoby o wyższym poziomie hamowania (przy kontroli wieku i tempa mówienia) są w stanie wykonać więcej korekt ukrytych niż osoby cechujące się niższym poziomem tej funkcji wykonawczej. Możliwe więc, że kontrola hamowania wcale nie wiąże się z częstszym i szybszym przerywaniem wypowiedzi (por. Engelhardta i in., 2013), ale właśnie z bardziej elastycznym dopasowaniem tempa mówienia do koniecznych przeformułowań w planie przedartykulacyjnym.

Uzyskane wyniki pokazują, że zapoczątkowany przed kilkunastu laty (por. Meyer i in., 2007; de Zubicaray i in., 2014) kierunek badań, poszukujących związków między kontrolą procesu mówienia, a bardziej ogólnymi mechanizmami kontroli poznawczej, przynosi interesujące rezultaty. Co więcej, dzięki lepszemu zrozumieniu tych związków możliwe staje się też wyjaśnienie postępującej kontroli procesu mówienia w odniesieniu do rozwojowych zmian w zakresie centralnej kontroli wykonawczej. Rezultaty podjętych w niniejszej pracy

analiz mogą stać się przyczynkiem do dalszych badań nad poznawczo-rozwojowymi aspektami kontroli procesu mówienia.

## Bibliografia

- Abdel Rahman, R., i Melinger, A. (2007). When bees hamper the production of honey: Lexical interference from associates in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 604–614.  
<https://doi.org/10.1037/0278-7393.33.3.604>
- Abel, S., Dressel, K., Kümmerer, D., Saur, D., Mader, I., Weiller, C., i Huber, W. (2009). Correct and erroneous picture naming responses in healthy subjects. *Neuroscience Letters*, 463(3), 167-171. <https://doi.org/10.1016/j.neulet.2009.07.077>
- Anderson, J. D., i Ofoe, L. C. (2019). The role of executive function in developmental stuttering. *Seminars in Speech and Language*, 40(4), 305-319. <https://doi.org/10.1055/s-0039-1692965>
- Arystoteles. (2004). Polityka. Wydawnictwo Naukowe PWN.
- Austin, J. L. (1962/1993). *Mówienie i poznawanie: Rozprawy i wykłady filozoficzne*. Wydawnictwo Naukowe PWN.
- Baddeley, A. D. (1998). *Pamięć: Poradnik użytkownika*. Prószyński i S-ka.
- Bartmiński, J. (2011). Droga naukowa Anny Wierzbickiej. Od składni polskiej prozy renesansowej do semantyki międzykulturowej. *Teksty Drugie. Teoria literatury, krytyka, interpretacja*, (1-2), 218-238.
- Baars, B. J. (1992). Preface. W: B. J. Baars (Red.), *Experimental slips and human error: Exploring the architecture of volition* (s. i-xix). Springer Science+Business Media.  
<https://doi.org/10.1007/978-1-4899-1164-3>
- Bajo, M. T., Puerta-Melguizo, M. C., i Macizo, P. (2003). The locus of semantic interference in picture naming. *Psicológica*, 24(1), 31–55.
- Beecher, M. D. (2021). Why are no animal communication systems simple languages? *Frontiers in Psychology*, 12, 701. <https://doi.org/10.3389/fpsyg.2021.602635>
- Berko Gleason, J. i Ratner, N. B. (Red.), *Psycholingwistyka*. Gdańskie Wydawnictwo Psychologiczne.
- Best, J. R., i Miller, P. H. (2010). A developmental perspective on executive function. *Child development*, 81(6), 1641–1660. <https://doi.org/10.1111/j.1467-8624.2010.01499.x>
- Bijak, E., i Kamińska, D. (2021). Kształtowanie płynności mówienia w terapii osób jękających się – perspektywa polska. W: K. Węsierska i H. Sønsterud (red.), *Dialog bez*

- barier – kompleksowa interwencja logopedyczna w jąkaniu (s. 207-265). Fundacja Wiedzy i Dialogu Społecznego Agere Aude.
- Blackmer, E. R., i Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition*, 39(1-2), 173–194.
- Bloom, I. (1980). Dlaczego nie gramatyka osiowa? [Reprint] W: G. W. Shugar i M. Smoczyńska (Red.), *Badania nad rozwojem języka dziecka*. Wybór prac (s. 240-255). Państwowe Wydawnictwo Naukowe.
- Bobrowski, I. (2009). O dwóch kognitywizmach. *Biuletyn Polskiego Towarzystwa Językoznawczego*, 65, 57-65.
- Bock, K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin & Review*, 3(4), 395-421.
- Bock, K., Dell, G. S., Garnsey, S. M., Kramer, A. F., i Kubose, T. T. (2007). Car talk, car listen. W: A. S. Meyer, L. R. Wheeldon, i A. Krott (Red.), *Automaticity and Control in Language Processing* (s. 21-42). Psychology Press.
- Boland, H. T., Hartsuiker, R. J., Pickering, M. J., i Postma, A. (2005). Repairing inappropriately specified utterances: Revision or restart? *Psychonomic Bulletin & Review*, 12(3), 472-477. <https://doi.org/10.3758/BF03193790>
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5(9/10), 341-345.
- Bowerman, M., i Choi, S. (2004). Kształtowanie znaczeń dla języka: Zjawiska uniwersalne i charakterystyczne dla danego języka w przyswajaniu kategorii semantycznych odnoszących się do przestrzeni [Reprint]. W B. Bokus i G. W. Shugar (Red.), *Psychologia języka dziecka* (s. 386-424). Gdańskie Wydawnictwo Psychologiczne.
- Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., i Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage*, 5(1), 49–62.
- Bręński, S. (2013). Błędy mowy w języku polskim i angielskim oraz czynniki wpływające na ich powstawanie. W: Łukasiewicz i E. J. Gorzelańczyk (Red.), *Psychologia języka* (s. 89-98). Wydawnictwo Uniwersytetu Kazimierza Wielkiego.
- Bręński, S. (2015). Rola reguł fonotaktycznych w powstawaniu przejęzyczeń. *Czasopismo Psychologiczne*, 21(2), 229-238.
- Bręński, S. (2020). Monitorowanie wypowiedzi a realizacja intencji w procesie mówienia. *Człowiek i Społeczeństwo*, 50, 47–64. <https://doi.org/10.14746/cis.2020.50.4>

- Brocklehurst, P. H. (2008). A review of evidence for the covert repair hypothesis of stuttering. *Contemporary Issues in Communication Science and Disorders*, 35, 25-43
- Brzezińska, A. I., i Nowotnik, A. (2012). Funkcje wykonawcze a funkcjonowanie dziecka w środowisku przedszkolnym i szkolnym. *Edukacja. Studia, badania, innowacje*, 1(117), 61-74.
- Budd, M.-J., Hanley, J. R., i Griffiths, Y. (2011). Simulating children's retrieval errors in picture-naming: A test of Foygel and Dell's (2000) semantic/phonological model of speech production. *Journal of Memory and Language*, 64(1), 74-87.  
<https://doi.org/10.1016/j.jml.2010.09.003>
- Bühler, K. (2004). Teoria języka: o językowej funkcji przedstawiania. Universitas
- Cholewiak, A. (2016). Metodologiczne aspekty badania tempa wypowiedzi (ujęcie neurolingwistyczne). *Annales Universitatis Paedagogicae Cracoviensis Studia Logopaedica*, 225(6), 1-10. <https://doi.org/10.24917/20837283.6.7>
- Cholin, J., Levelt, W. J., i Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition*, 99(2), 205-235. <https://doi.org/10.1016/j.cognition.2005.01.009>
- Chomsky, N. (1982). *Zagadnienia teorii składni*. Zakład Narodowy im. Ossolińskich.
- Chuderski, A. (2010). Samokontrola: własności, funkcje, mechanizmy i ograniczenia. *Studia z Kognitywistyki i Filozofii Umysłu*, 4(1), 27–51.
- Clark, E. V. (1978). Awareness of Language: Some Evidence from what Children Say and Do. W: A. Sinclair, R. J. Jarvella, i W. J. M. Levelt (Red.), *The Child's Conception of Language* (s. 17-43). Springer.
- Clark, H. H. (1996). *Using language*. Cambridge University Press
- Cohen, J., Cohen, P., West, S. G., i Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates
- Conover, W. J. (1999). *Practical Nonparametric Statistics* (3rd ed.). John Wiley & Sons, Inc.
- Crosbie, S., Holm, A., i Dodd, B. (2009). Cognitive flexibility in children with and without speech disorder. *Child Language Teaching and Therapy*, 25(2), 250-270.  
<https://doi.org/10.1177/0265659009102990>
- Cutler, A. (2006). Rudolf Meringer. W: K. Brown (Red.), *Encyclopedia of Language and Linguistics* (vol. 8) (s. 12-13). Elsevier.

- Cutler, A., i Fay, D. (Red.). (1978). Introduction. W: A. Cutler i D. Fay (Red.), *Versprechen und Verlesen: Ein psychologisch-linguistischer Essay* [Ponowne wydanie z komentarzami R. Meringer i C. Mayer, 1895] (s. ix-xl). John Benjamins.
- De Boysson-Bardies, B., Halle, P., Sagart, L., i Durand, C. (1989). A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*, 16, 1-17.
- De Zubicaray, G. I., Hartsuiker, R. J., i Acheson, D. J. (2014). Mind what you say—general and specific mechanisms for monitoring in speech production. *Frontiers in Human Neuroscience*, 8, 514. <https://doi.org/10.3389/fnhum.2014.00514>
- De Saussure, F. (1961). *Kurs językoznawstwa ogólnego*. Państwowe Wydawnictwo Naukowe.
- De Saussure, F. (1991). *Kurs językoznawstwa ogólnego* (Wyd. 2 popr.). Państwowe Wydawnictwo Naukowe.
- Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain*. W. W. Norton & Company.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283–321.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., i Gagnon, D. A. (1997). Lexical Access in Aphasic and Nonaphasic Speakers. *Psychological Review*, 104(4), 801-838
- Dell, G. S., Oppenheim, G. M., i Kittredge, A. K. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language and Cognitive Processes*, 23(4), 583–608.
- Dhooge, E., i Hartsuiker, R. J. (2010). The distractor frequency effect in picture–word interference: Evidence for response exclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(4), 878–891. <https://doi.org/10.1037/a0019128>.
- Dhooge, E., i Hartsuiker, R. J. (2013). Distractor exclusion is not an early process: A reply to Roelofs, Piai, and Schriefers (2011). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(1), 313–316. <https://doi.org/10.1037/a0028473>
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Dunaj, M. (2015). *GLUCHY-ŚWIAT. Gluchota w perspektywie antropologii zaangażowanej*. Wydział Filozoiczno-Historyczny Uniwersytetu Łódzkiego.
- Eco, U. (2002). *W poszukiwaniu języka uniwersalnego*. Wydawnictwo Marabut, Oficyna Wydawnicza Volumen.

- Eggers, K., De Nil, L. F., i Van den Bergh, B. R. H. (2013). Inhibitory control in childhood stuttering. *Journal of Fluency Disorders*, 38(1), 1-13
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., i Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303-306.
- Engelhardt, P. E., Nigg, J. T., i Ferreira, F. (2013). Is the fluency of language outputs related to individual differences in intelligence and executive function? *Acta Psychologica*, 144, 424–432.
- Eriksen, B. A., i Eriksen, C. W. (1974). Effects of noise letters upon the identification of target letters in a non-search task. *Perception & Psychophysics*, 16(1), 143–149
- Evans, M. A. (1985). Self-Initiated Speech Repairs: A Reflection of Communicative Monitoring in Young Children. *Developmental Psychology*, 21(2), 365-371.
- Everett, D. (2005). Cultural constraints on grammar and cognition in Piraha: Another look at the design features of human language. *Current Anthropology*, 46(4), 621-646.
- Fay, D., i Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry*, 8(3), 505-520
- Fąfrowicz, M., i Marek, T. (2008). Przedni zakręt kory obręczy – perspektywa neurokogniwiwistyczna. *Przegląd Psychologiczny*, 51(2), 149-160.
- Ferguson, C. A., i Farwell, C. B. (1975). Words and Sounds in Early Language Acquisition. *Language*, 51(2), 419-439.
- Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532–538.  
<https://doi.org/10.1037/a0015808>
- Fernyhough, C. (2004). Alien voices and inner dialogue: Towards a developmental account of auditory verbal hallucinations. *New Ideas in Psychology*, 22(1), 49–68
- Field, A. (2009). *Discovering Statistics Using SPSS* (3rd ed.). Sage Publications.
- Fisiak, J. (1985). *Wstęp do współczesnych teorii lingwistycznych*. Wydawnictwa Szkolne i Pedagogiczne.
- Fisk, J. E., i Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology*, 26(7), 874-890. <https://doi.org/10.1080/13803390490510680>

- Forster, K. I., i Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers*, 35(1), 116–124.  
<https://doi.org/10.3758/BF03195503>
- Forrester, M. A. (2008). The Emergence Of Self-Repair: A Case Study Of One Child During The Early Preschool Years. *Research on Language and Social Interaction*, 41(1), 99-128.  
<https://doi.org/10.1080/08351810701691206>
- Fox, C. J., Mueller, S. T., Gray, H. M., Raber, J., i Piper, B. J. (2013). Evaluation of a short-form of the Berg Card Sorting Test. *PLoS ONE*, 8(5), e63885.  
<https://doi.org/10.1371/journal.pone.0063885>
- Franken, M. K., Liu, B. C., i Ostry, D. J. (2022). Towards a somatosensory theory of speech perception. *Journal of Neurophysiology*, 128(12), 1683–1695.  
<https://doi.org/10.1152/jn.00381.2022>
- Fromkin, V. A. (1973). The Non-Anomalous Nature of Anomalous Utterances. W: V. A. Fromkin (Red.), *Speech Errors as Linguistic Evidence* (s. 215-242). Mouton.
- Fromkin, V. A., i Ratner, N. B. (2005). Wytwarzanie mowy. W: J. B. Gleason i N. B. Ratner (Red.), *Psycholingwistyka* (s. 333–374). Gdańskie Wydawnictwo Psychologiczne.
- Fry, D. B. (1969). The linguistic evidence of speech errors. *Brno Studies in English*, 8, 69-74.
- Frydrychowicz, S. (1988). Język w rozwoju jednostkowym. W: T. Rzepa i S. Frydrychowicz (Red.), *Relacje między doświadczeniem językowym a doświadczeniem indywidualnym* (s. 54-65). Zakład Narodowy im. Ossolińskich.
- Frydrychowicz, S. (1999). *Proces mówienia: Wybrane psychologiczne aspekty na przykładzie interpretacji zdania niejednoznacznego*. Wydawnictwo Naukowe UAM.
- Gage, N. M., i Baars, B. J. (2013). *Fundamentals of cognitive neuroscience: A beginner's guide* (2nd ed.). Elsevier Academic Press
- Garrett, M.F. (1976). Syntactic processes in sentence production. W: R.J. Wales i E.C.T. Walker (Red.), *New approaches to language mechanisms* (s. 231–255). North Holland.
- Garrett, M. F. (1980). The limits of accommodation. W: V. Fromkin (Red.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (s. 263-271). Academic Press.
- Garrett, M. F. (1993). The Organization of Processing Structure for Language Production: Applications to Aphasic Speech. W: P. R. Thagard (Red.), *Computational Philosophy of Science*. The MIT Press.

- Gauvin, H. S., De Baene, W., Brass, M., i Hartsuiker, R. J. (2016). Conflict monitoring in speech processing: An fMRI study of error detection in speech production and perception. *NeuroImage*, 126, 96-105. <https://doi.org/10.1016/j.neuroimage.2015.11.037>
- Geva, S., i Fernyhough, C. (2019). A penny for your thoughts: Children's inner speech and its neuro-development. *Frontiers in Psychology*, 10, Article 1708. <https://doi.org/10.3389/fpsyg.2019.01708>
- George, D., i Mallery, P. (2019). *IBM SPSS Statistics 26 step by step: A simple guide and reference (16th ed.)*. Routledge.
- Glaser, M. O., i Dünghoff, F.-J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception and Performance*, 10(5), 640-654.
- Goddard, C. (2012). Semantic primes, semantic molecules, semantic templates: Key concepts in the NSM approach to lexical typology. *Linguistics*, 50(3), 711-743.
- Gołąb, Z., Heinz, A., i Polański, K. (1970). *Słownik terminologii językoznawczej*. Państwowe Wydawnictwo Naukowe.
- Grice, P. (1980). Logika a konwersacja. W: S. Barbara (Red.), *Język w świetle nauki* (s. 91–114). Wydawnictwo Czytelnik.
- Grzegorzczkova, R. (2001). *Wprowadzenie do semantyki językoznawczej*. Wydawnictwo Naukowe PWN.
- Grzegorzczkova, R. (2007). *Wstęp do językoznawstwa*. Wydawnictwo Naukowe PWN.
- Hanley, J. R., Cortis, C., Budd, M. J., i Nozari, N. (2016). Did I say dog or cat? A study of semantic error detection and correction in children. *Journal of Experimental Child Psychology*, 142, 36–47. <https://doi.org/10.1016/j.jecp.2015.09.008>
- Harley, T. A. (1984). A Critique of Top-down Independent Levels Models of Speech Production: Evidence from Non-plan-Internal Speech Errors. *Cognitive Science*, 8(3), 191-219. [https://doi.org/10.1207/s15516709cog0803\\_1](https://doi.org/10.1207/s15516709cog0803_1)
- Harley, T. A. (2001). *The Psychology of Language: From Data to Theory*. Taylor & Francis.
- Harley, T. A., i MacAndrew, S. B. G. (2001). Constraints upon word substitution speech errors. *Journal of psycholinguistic research*, 30(4), 395–418. <https://doi.org/10.1023/a:1010421724343>

- Hartsuiker, R. J. (2007). Studies on verbal self-monitoring: The perceptual loop model and beyond. W: A. S. Meyer, L. R. Wheeldon, i A. Krott (Red.), *Automaticity and Control in Language Processing* (s. 93–121). Psychology Press.
- Hartsuiker, R. J., i Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology*, 42(2), 113–157. <https://doi.org/10.1006/cogp.2000.0744>
- Hartsuiker, R. J., Bastiaanse, R., Postma, A., i Wijnen, F. (2005). Phonological encoding and monitoring in normal and pathological speech. W: R. J. Hartsuiker, R. Bastiaanse, A. Postma, i F. Wijnen (Red.), *Phonological Encoding and Monitoring in Normal and Pathological Speech* (s. 1-14). Psychology Press.
- Hartsuiker, R. J., Kolk, H. H. J., i Lickley, R. J. (2005). Stuttering on function words and content words: A computational test of the covert repair hypothesis. W: R. J. Hartsuiker, R. Bastiaanse, A. Postma, i F. Wijnen (Red.), *Phonological Encoding and Monitoring in Normal and Pathological Speech* (s. 261-280). Psychology Press.
- Hartsuiker, R. J., Kolk, H. H. J., i Martensen, H. (2005). The division of labor between internal and external speech monitoring. W: R. J. Hartsuiker, R. Bastiaanse, A. Postma, i F. Wijnen (Red.), *Phonological Encoding and Monitoring in Normal and Pathological Speech* (s. 189-205). Psychology Press.
- Hartsuiker, R. J., i Notebaert, L. (2010). Lexical access problems lead to disfluencies in speech. *Experimental Psychology*, 57(3), 169–177. <https://doi.org/10.1027/1618-3169/a000021>.
- Hartsuiker, R. J., i Moors, A. (2017). On the automaticity of language processing. W: H.-J. Schmid (Red.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (s. 201–225). American Psychological Association; De Gruyter Mouton. <https://doi.org/10.1037/15969-010>
- Harvey, D. Y., Traut, H. J., i Middleton, E. L. (2019). Semantic interference in speech error production in a randomized continuous naming task: Evidence from aphasia. *Language, cognition and neuroscience*, 34(1), 69–86. <https://doi.org/10.1080/23273798.2018.1501500>
- Higby, E., Cahana-Amitay, D., Vogel-Eyny, A., Spiro, A. III, Albert, M. L., i Obler, L. K. (2019). The Role of Executive Functions in Object- and Action-Naming among Older Adults. *Experimental Aging Research*, 45(4), 306–330

- Hjelmslev, L. (1979). Prolegomena do teorii języka. W: H. Kurkowska i A. Weinsberg (Red.), *Językoznawstwo strukturalne. Wybór tekstów* (s. [numery stron]). Państwowe Wydawnictwo Naukowe.
- Hockett, C. F. (1968). *Kurs językoznawstwa współczesnego*. Państwowe Wydawnictwo Naukowe.
- Hockett, C. F. (1960). The Origin of Speech. *Scientific American*, 203(1), 88–111.
- Hockett, C. F. (1979). Zagadnienie uniwersaliów w języku. W: H. Kurkowska i A. Weinsberg (Red.), *Językoznawstwo strukturalne: wybór tekstów* (s. 209-229). Państwowe Wydawnictwo Naukowe.
- Hollander, M., Wolfe, D. A., i Chicken, E. (2014). *Nonparametric Statistical Methods*. John Wiley & Sons.
- Hollister, J., Van Horne, A. O., i Zebrowski, P. (2017). The Relationship Between Grammatical Development and Disfluencies in Preschool Children Who Stutter and Those Who Recover. *American Journal of Speech-Language Pathology*, 26, 44-56.  
[https://doi.org/10.1044/2016\\_AJSLP-15-0022](https://doi.org/10.1044/2016_AJSLP-15-0022)
- Hornowska, E., Brzezińska, A., Appelt, K., i Kaliszewska-Czeremska, K., (2014). *Rola środowiska w rozwoju małego dziecka – metody badania*. Wydawnictwo Naukowe Scholar.
- Huettig, F., i Hartsuiker, R. J. (2010). Listening to yourself is like listening to others: External, but not internal, verbal self-monitoring is based on speech perception. *Language and Cognitive Processes*, 25(3), 347–374. <https://doi.org/10.1080/01690960903046926>
- Huizinga, M., Dolan, C. V., i van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, 44(11), 2017-2036. <https://doi.org/10.1016/j.neuropsychologia.2006.01.010>
- Humphreys, G. W., Forde, E. M. E., Steer, E., Samson, D., i Connelly, C. (2007). Executive functions in name retrieval: Evidence from neuropsychology. W: A. S. Meyer, L. R. Wheeldon, i A. Krott (Red.), *Automaticity and Control in Language Processing* (s. 143-159). Psychology Press.
- Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, 2, 255. <https://doi.org/10.3389/fpsyg.2011.00255>
- Indefrey, P., i Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144

- Ingram, D. (1986). Phonological development: production. W: P. Fletcher i M. Garman (Red.), *Language Acquisition Studies in First Language Development* (pp. 223-239). Cambridge University Press. <https://doi.org/10.1017/CBO9780511620683.013>
- Izert, M., i Pachocińska, E. (1998). *Wstęp do językoznawstwa ogólnego*. Uniwersytet Warszawski - Instytut Romanistyki.
- Jankowski, P. (2012). Funkcje wykonawcze a osiągnięcia dzieci w wieku wczesnoszkolnym. *Edukacja*, 1(117), 75-86.
- Jaeger, J. J. (2005). *KIDS' SLIPS: What Young Children's Slips of the Tongue Reveal About Language Development*. Lawrence Erlbaum Associates.
- Jerger, S., Martin, R. C., i Damian, M. F. (2002). Semantic and phonological influences on picture naming by children and teenagers. *Journal of Memory and Language*, 47(2), 229–249. [https://doi.org/10.1016/S0749-596X\(02\)00002-5](https://doi.org/10.1016/S0749-596X(02)00002-5)
- Jescheniak, J. D., i Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 824-843.
- Jescheniak, J. D., Wöhner, S., Bethcke, H. S., i Beaupain, M. C. (2020). Semantic interference in the picture-word interference task: Is there a pre-lexical, conceptual contribution to the effect? *Psychonomic Bulletin & Review*, 27, 373–378. <https://doi.org/10.3758/s13423-019-01667-w>
- Jodzio, K. (2008). *Neuropsychologia intencjonalnego działania: Koncepcje funkcji wykonawczych*. Warszawa: Wydawnictwo Naukowe Scholar.
- Jodzio, K., Biechowska, D., Szurowska, E., i Gąsecki, D. (2012). Profilowa analiza dysfunkcji wykonawczych w diagnostyce neuropsychologicznej osób po udarze mózgu. *Roczniki Psychologiczne*, 15(3), 83-100.
- Józefacka-Szram, N. M. (2014). Diagnoza funkcji wykonawczych u dzieci. *Psychiatria i Psychologia Kliniczna*, 14(2), 116–121. <https://doi.org/10.15557/PiPK.2014.0016>
- Juszyk, P. W., Friederici, A. D., Wessels, J., Svenkerud, V. Y., i Juszyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32, 402-420.
- Kaczmarek, L. (1977). *Nasze dziecko uczy się mowy*. Wydawnictwo Lubelskie.
- Kalisz, R. (2001). *Językoznawstwo kognitywne w świetle językoznawstwa funkcjonalnego*. Wydawnictwo Uniwersytetu Gdańskiego.

- Kawachi, K. (2002). Practice effects on speech production planning: Evidence from slips of the tongue in spontaneous vs. preplanned speech in Japanese. *Journal of Psycholinguistic Research*, 31(4), 363-390. doi:10.1023/A:1019569724949
- Krasowicz-Kupis, G. (2004). *Rozwój świadomości językowej dziecka. Teoria i praktyka*. Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.
- Króliczak, G., Buchwald, M., Kleka, P., Klichowski, M., Potok, W., Nowik, A. M., Randerath, J., i Piper, B. J. (2021). Manual praxis and language-production networks, and their links to handedness. *Cortex*, 140, 110–127
- Kujałowicz, A., i Zajdler, E. (2009). Language activation and lexical processing by Polish learners of English and Chinese—the role of learning experience in TLA. *International Journal of Multilingualism*, 6(1), 85-104.
- Kurcz, I. (1987). *Język a reprezentacja świata w umyśle*. Państwowe Wydawnictwo Naukowe.
- Kurcz, I. (2005). *Psychologia języka i komunikacji*. Wydawnictwo Naukowe Scholar.
- Kurcz, I. (2011). Charakterystyka kompetencji językowej - wytwarzanie i percepcja mowy. W I. Kurcz & H. Okuniewska (Red.), *Język jako przedmiot badań psychologicznych: Psycholingwistyka ogólna i neurolingwistyka* (s. 66-79). Wydawnictwo SWPS Academica.
- Kurcz, I., Lewicki, A., Sambor, J., Szafran, K., i Woronczak, J. (1990). *Słownik frekwencyjny polszczyzny współczesnej* (t. 1-2). Polska Akademia Nauk Instytut Języka Polskiego.
- Kwarciak, B. (1995). *Początki i podstawowe mechanizmy świadomości metajęzykowej*. Wydawnictwo Uniwersytetu Jagiellońskiego.
- Laakso, M. (2010). Children's emerging and developing self-repair practices. W: H. Gardner i M. Forrester (Red.), *Analysing Interactions in Childhood: Insights from Conversation Analysis* (s. 74-99). Wiley-Blackwell.
- Ladd, D. R. (2012). What is duality of patterning, anyway? *Language and Cognition*, 4(4), 261–273. doi:10.1515/langcog-2012-0015
- Lambon Ralph, M. A., i Fillingham, J. K. (2007). The importance of memory and executive function in aphasia: Evidence from the treatment of anomia using errorless and errorful learning. W: A. S. Meyer, L. R. Wheeldon, i A. Krott (Red.), *Automaticity and Control in Language Processing* (s. 193-216). Psychology Press
- Langacker, R. W. (2009). *Gramatyka kognitywna: Wprowadzenie*. Towarzystwo Autorów i Wydawców Prac Naukowych Universitas.

- Lehto, J. E., Juujärvi, P., Kooistra, L., i Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, 21, 59-80.
- Levitt, A. G., i Utman, J. G. (1992). From babbling towards the sound systems of English and French: A longitudinal two-case study. *Journal of Child Language*, 19, 19-49.
- Levy, Y. (1999). Early metalinguistic competence: Speech monitoring and repair behavior. *Developmental Psychology*, 35(3), 822–834. <https://doi.org/10.1037/0012-1649.35.3.822>
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14(1), 41–104. [https://doi.org/10.1016/0010-0277\(83\)90026-4](https://doi.org/10.1016/0010-0277(83)90026-4)
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. The MIT Press.
- Levelt, W. J. M. (1992). The perceptual loop theory not disconfirmed: A reply to MacKay. *Consciousness and Cognition*, 1(3), 226-230. [https://doi.org/10.1016/1053-8100\(92\)90062-F](https://doi.org/10.1016/1053-8100(92)90062-F)
- Levelt, W. J. M. (1993). Lexical Access in Speech Production. W: E. Reuland i W. Abraham (Red.), *Knowledge and Language* (s. 241–251). Springer.
- Levelt, W. J. M. (1998). The genetic perspective in psycholinguistics, or: Where do spoken words come from? *Journal of Psycholinguistic Research*, 27(2), 167-180. doi:10.1023/A:1023245931630.
- Levelt, W. J. M. (1999a). Models of word production. *Trends in Cognitive Sciences*, 3, 223-232.
- Levelt, W. J. M. (1999b). Producing spoken language: A blueprint of the speaker. W: C. M. Brown, & P. Hagoort (Red.), *The neurocognition of language* (s. 83-122). Oxford University Press.
- Levelt, W. J. M. (2000). Speech production. W: A. E. Kazdin (Red.), *Encyclopedia of psychology* (s. 432–433). Oxford University Press.
- Levelt, W. J. M. (2001a). Relations between speech production and speech perception: Some behavioral and neurological observations. W: E. Dupoux (Red.), *Language, brain and cognitive development: Essays in honour of Jacques Mehler* (s. 241-256). MIT Press.
- Levelt, W. J. M. (2001b). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, 98(22), 13464-13471. <https://doi.org/10.1073/pnas.231459498>
- Levelt, W. J. M., Roelofs, A., i Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1–75.

- Lind, A., Hall, L., Breidegard, B., Balkenius, C., i Johansson, P. (2014). Speakers' Acceptance of Real-Time Speech Exchange Indicates That We Use Auditory Feedback to Specify the Meaning of What We Say. *Psychological Science*, 25(6), 1198-1205.  
<https://doi.org/10.1177/0956797614529797>
- Lind, A., i Hartsuiker, R. J. (2020). Self-Monitoring in Speech Production: Comprehending the Conflict Between Conflict- and Comprehension-Based Accounts. *Journal of Cognition*, 3(1), 16, 1-4. <https://doi.org/10.5334/joc.118>
- Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception & Psychophysics*, 39(3), 155–158. doi: 10.3758/bf03212485
- Luna, B. (2009). Developmental changes in cognitive control through adolescence. *Advances in Child Development and Behavior*, 37, 233-278
- Lyons, J. (1984). *Semantyka I*. Państwowe Wydawnictwo Naukowe.
- Lyons, J. (1998). *Chomsky*. Warszawa: Prószyński i S-ka.
- Łozińska, S., i Rutkowski, P. (2014). Ikoniczność w polskim języku migowym. W: P. Rutkowski i S. Łozińska (Red.), *Lingwistyka przestrzeni i ruchu. Komunikacja migowa a metody korpusowe* (s. 132–144). Wydział Polonistyki Uniwersytetu Warszawskiego.
- Łobacz, P. (2005). Prawidłowy rozwój mowy dziecka. W: T. Gałkowski, E. Szelaż, i G. Jastrzębowska (Red.), *Podstawy neurologopedii. Podręcznik akademicki* (s. 231-268). Wydawnictwo Uniwersytetu Opolskiego.
- Łuczyński, E. (2019). Proces gramatyzacji w rozwoju językowym dziecka. W: M. Kielar-Turska i S. Milewski (Red.), *Język w biegu życia* (s. 171-201). Harmonia Universalis.
- MacDonald, E. N., Johnson, E. K., Forsythe, J., Plante, P., i Munhall, K. G. (2012). Children's development of self-regulation in speech production. *Current Biology*, 22(1), 113-117.  
<https://doi.org/10.1016/j.cub.2011.11.052>
- MacKay, D. G. (1980). Speech errors: retro spect and prospect. W: V. Fromkin (Red.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (s. 319-332). Academic Press.
- MacKay, D. G. (1992). Errors, ambiguity, and awareness in language perception and production. W: B. J. Baars (Red.), *Experimental slips and human error: Exploring the architecture of volition* (str. 39-69). Springer Science+Business Media.  
[https://doi.org/10.1007/978-1-4899-1164-3\\_2](https://doi.org/10.1007/978-1-4899-1164-3_2)
- MacNeilage, P. F. (2005). *The origin of speech*. Oxford University Press.

- Mandler, J. M. (1992). How to Build a Baby: II. Conceptual Primitives. *Psychological Review*, 99(4), 587-604.
- Manfra, L., Tyler, S. L., i Winsler, A. (2016). Speech monitoring and repairs in preschool children's social and private speech. *Early Childhood Research Quarterly*, 37, 94-105. <https://doi.org/10.1016/j.ecresq.2016.04.004>
- Maruszewski, M. (1970). *Mowa a mózg: Zagadnienia neuropsychologiczne*. Państwowe Wydawnictwo Naukowe.
- Mattson, M. E., i Baars, B. J. (1992). Error-minimizing mechanisms: Boosting or editing? W B. J. Baars (Red.), *Experimental slips and human error: Exploring the architecture of volition* (str. 263-287). Springer Science+Business Media. [https://doi.org/10.1007/978-1-4899-1164-3\\_11](https://doi.org/10.1007/978-1-4899-1164-3_11)
- Maycock, K. W., Lambert, J., i Bane, D. (2018). Flipping learning not just content: A 4-year action research study investigating the appropriate level of flipped learning. *Journal of Computer Assisted Learning*, 34(5), 491-505. <https://doi.org/10.1111/jcal.12274>
- McAuley, T., i White, D. (2011). A latent variables examination of processing speed, response inhibition, and working memory during typical development. *Journal of Experimental Child Psychology*, 108(3), 453-468. <https://doi.org/10.1016/j.jecp.2010.08.009>
- Mecner, P. (2007). *Elementy gramatyki umysłu: Od struktur składniowych do minimalizmu*. Towarzystwo Autorów i Wydawców Prac Naukowych Universitas
- Meyer, A. S., Wheeldon, L. R., i Krott, A. (2007). Preface. W: A. S. Meyer, L. R. Wheeldon, i A. Krott (Red.), *Automaticity and Control in Language Processing* (str. ix-xiii). Psychology Press.
- Michalik, M., Milewski, S., Kaczorowska-Bray, K., i Solak, A. (2018). Tempo wypowiedzi jako kategoria badawcza – w normie i w zaburzeniach. *Annales Universitatis Paedagogicae Cracoviensis. Studia Linguistica*, (13), 107–116. <https://doi.org/10.24917/20831765.13.8>
- Miles, S., Howlett, C. A., Berryman, C., Nedeljkovic, M., Moseley, G. L., i Phillipou, A. (2021). Considerations for using the Wisconsin Card Sorting Test to assess cognitive flexibility. *Behavior Research Methods*, 53(5), 2083–2091. <https://doi.org/10.3758/s13428-021-01551-3>
- Milewski, T. (1967/2004). *Językoznawstwo*. Wydawnictwo Naukowe PWN.
- Miller, G. A., Galanter, E., i Pribram, K. H. (1980). *Plany i struktura zachowania*. Państwowe Wydawnictwo Naukowe.

- Miłkowski, M. (2007). Architektury modularne: Funkcje, indywidualizacja modułów i poziomy opis. W: S. Wróbel (Red.), *Modularność umysłu* (s. 53-76). Wydział Pedagogiczno Artystyczny UAM.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., i Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100.  
<https://doi.org/10.1006/cogp.1999.0734>
- Morgan, J. L., i Wheeldon, L. R. (2003). Syllable monitoring in internally and externally generated English words. *Journal of Psycholinguistic Research*, 32(3), 269-296.
- Motley, M. T. (1980). Verification of ‘Freudian slips’ and semantic prearticulatory editing via laboratory-induced spoonerisms. W: V. A. Fromkin (Red.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (s. 133-147). Academic Press.
- Motley, M. T., Camden, C. T., i Baars, B. J. (1982). Covert formulation and editing of anomalies in speech production: Evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, 21(5), 578–594.  
[https://doi.org/10.1016/S0022-5371\(82\)90791-5](https://doi.org/10.1016/S0022-5371(82)90791-5)
- Mueller, S. T., i Piper, B. J. (2014). The PEBL Psychological Experiment Building Language PEBL and PEBL Test Battery. *Journal of Neuroscience Methods*, 222, 250-259.  
<https://doi.org/10.1016/j.jneumeth.2013.10.024>
- Newman, R. S., i German, D. J. (2005). Life span effects of lexical factors on oral naming. *Language and Speech*, 48(2), 123–156. <https://doi.org/10.1177/00238309050480020101>
- Nęcka, E. (2003). *Inteligencja: Geneza, struktura, funkcje*. Gdańskie Wydawnictwo Psychologiczne.
- Nęcka, E., Orzechowski, J., i Szymura, B. (2006). *Psychologia poznawcza*. Wydawnictwo Academica SWSP i Wydawnictwo Naukowe PWN.
- Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., i Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, 12(2), 241-268.  
<https://doi.org/10.3758/s13415-011-0083-5>
- Niewiarowski, J., Mroziński, B., i Morawiak, A. (2013). Jednoczynnikowa analiza wariancji z powtarzonym pomiarem. W: S. Bedyńska, i M. Cypryńska (Red.), *Statystyczny drogowskaz 2. Praktyczne wprowadzenie do analizy wariancji* (s. 113-133). Warszawa: Wydawnictwo Akademickie Sedno.

- Nigg, J.T. (2000). On inhibition/disinhibition in developmental psychology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, 126(2), 220–246
- Nooteboom, S. G. (1980). Speaking and unspeaking: Detection and correction of phonological and lexical errors in spontaneous speech. W: V. A. Fromkin (Red.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (s. 87-95). Academic Press.
- Nooteboom, S. G., i Quené, H. (2019). Temporal aspects of self-monitoring for speech errors. *Journal of Memory and Language*, 105(1), 43-59.  
<https://doi.org/10.1016/j.jml.2018.11.002>
- Nozari, N. (2018). How special is language production? Perspectives from monitoring and control. In K. D. Federmeier i D. G. Watson (Red.), *The psychology of learning and motivation: Current topics in language* (s. 179–213). Elsevier Academic Press.
- Nozari, N. (2020). A Comprehension- or a Production-Based Monitor? Response to Roelofs (2020). *Journal of Cognition*, 3(1), 19, 1-21. <https://doi.org/10.5334/joc.102>
- Nozari, N., Dell, G. S., i Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology*, 63(1), 1–33. <https://doi.org/10.1016/j.cogpsych.2011.05.001>
- Ofoe, L. C., Anderson, J. D., i Ntourou, K. (2018). Short-term memory, inhibition, and attention in developmental stuttering: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 61(7), 1626-1648. [https://doi.org/10.1044/2018\\_JSLHR-S-17-0372](https://doi.org/10.1044/2018_JSLHR-S-17-0372)
- Oomen, C. C. E., Postma, A., i Kolk, H. H. J. (2005). Speech monitoring in aphasia: Error detection and repair behaviour in a patient with Broca's aphasia. W: R. J. Hartsuiker, R. Bastiaanse, A. Postma, i F. Wijnen (Red.), *Phonological Encoding and Monitoring in Normal and Pathological Speech* (s. 209-225). Psychology Press.
- Orzechowski, J. (2012). *Magiczna liczba jeden: Czyli co jeszcze zmieści się w pamięci roboczej*. Wydawnictwo Uniwersytetu Jagiellońskiego.
- Özdemir, R., Roelofs, A., i Levelt, W. J. M. (2007). Perceptual uniqueness point effects in monitoring internal speech. *Cognition*, 105(2), 457-465.  
<https://doi.org/10.1016/j.cognition.2006.10.006>
- Paphiti, M., i Eggers, K. (2022). Cognitive flexibility in younger and older children who stutter. *Frontiers in Psychology*, 13, Article 1017319.  
<https://doi.org/10.3389/fpsyg.2022.1017319>

- Pąchalska, M., Kaczmarek, B. L. J., i Kropotov, J. D. (2014). *Neuropsychologia kliniczna: Od teorii do praktyki*. Wydawnictwo Naukowe PWN.
- Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., i Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *Journal of Experimental Child Psychology*, 108(3), 621-637.  
<https://doi.org/10.1016/j.jecp.2010.09.008>
- Peterson, E., i Welsh, M.C. (2014). The development of hot and cool executive functions in childhood and adolescence: Are we getting warmer? W: S. Goldstein i J.A. Naglieri (Red.), *Handbook of Executive Functioning* (s. 45–65). Springer Science + Business Media.  
[https://doi.org/10.1007/978-1-4614-8106-5\\_4](https://doi.org/10.1007/978-1-4614-8106-5_4)
- Piaget, J. (1966). *Studia z psychologii dziecka*. Państwowe Wydawnictwo Naukowe.
- Piaget, J., i Inhelder, B. (1996). *Psychologia dziecka*. Siedmioróg.
- Piai, V., Roelofs, A., Jensen, O., Schoffelen, J.-M., i Bonnefond, M. (2014). Distinct patterns of brain activity characterise lexical activation and competition in spoken word production. *PLoS ONE*, 9(2), e88674. <https://doi.org/10.1371/journal.pone.0088674>
- Pillai, S. (2006). Self-monitoring and self-repair in spontaneous speech. *K@ta: A Biannual Publication on the Study of Language and Literature*, 8(2), 114–126.  
<https://doi.org/10.9744/kata.8.2.114-126>
- Piotrowski, K. T., Stettner, Z., Orzechowski, J., i Balas, R. (2009). Jak działa pamięć robocza? W: J. Orzechowski, K. T. Piotrowski, R. Balas, i Z. Stettner (Red.), *Pamięć robocza* (s. 25-46). Wtdawnictwo SWPS Academica.
- Piotrowski, K.T., Stettner, Z., Wierzchoń, M., Balas, R., i Bielecki, M. (2009). Eksperymentalne techniki badania pamięci roboczej. W: J. Orzechowski, K.T. Piotrowski, R. Balas, i Z. Stettner (Red.), *Pamięć robocza* (s. 47-76). Wydawnictwo SWPS Academica.
- Polański, K. (1971). Gramatyka transformacyjno-generatywna a inne teorie języka. *Biuletyn Polskiego Towarzystwa Językoznawczego*, (XXVIII), 87–98.
- Polański, K. (Red.) (2003). *Encyklopedia językoznawstwa ogólnego*. Wydawnictwo Zakładu Narodowego imienia Ossolińskich.
- Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, 77(2), 97–131
- Postma, A., i Kolk, H. (1993). The Covert Repair Hypothesis: Prearticulatory Repair Processes in Normal and Stuttered Disfluencies. *Journal of Speech and Hearing Research*, 36, 472-487.

- Postma, A., i Oomen, C. C. E. (2005). Critical issues in speech monitoring. W: R. J. Hartsuiker, R. Bastiaanse, A. Postma i F. Wijnen (Red.), *Phonological Encoding and Monitoring in Normal and Pathological Speech* (s. 157-166). Psychology Press.
- Prechtel, P. (2007). *Wprowadzenie do filozofii języka*. Wydawnictwo WAM.
- Puppel, S. (1994). Some aspects of the monitoring (regulation) function in speech production. *Studia Anglica Posnaniensia*, 28, 72-77.
- Puppel, S. (2006). Czarna Skrzynka Noama Chomsky'ego a trychotomiczny model umysłu: Perspektywa psycholingwistyczna. *Electronic Journal Oikeios Logos*, (2), 1-21.
- Putko, A. (2008). *Dziecięca 'teoria umysłu' w fazie jawnej i utajonej a funkcje wykonawcze*. Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza.
- Putko, A. (2011). Teoria umysłu a zimne versus gorące funkcje zarządzające u dzieci w wieku przedszkolnym. *Psychologia Rozwojowa*, 16(1), 73–84.
- Rajtar, A. M., Przewoźnik, D. A., Starowicz-Filip, A., i Będkowska-Korpała, B. (2014). Dysfunkcje wykonawcze w afatycznych zaburzeniach mowy po udarze mózgu. *Postępy Psychiatrii i Neurologii*, 23(1), 41-46. <https://doi.org/10.1016/j.pin.2014.03.004>
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, 6(2), 135-147. <https://doi.org/10.1016/j.edurev.2010.12.001>
- Ridderinkhof, K. R., Wylie, S. A., van den Wildenberg, W. P. M., Bashore, T. R., Jr., i van der Molen, M. W. (2021). The arrow of time: Advancing insights into action control from the arrow version of the Eriksen flanker task. *Attention, Perception, & Psychophysics*, 83(2), 700–721
- Rispoli, M. (2003). Changes in the nature of sentence production during the period of grammatical development. *Journal of Speech, Language, and Hearing Research*, 46(4), 818-830.
- Reason, J. (1984). Lapses of Attention in Everyday Life. W: R. Parasuraman i D. R. Davies (Red.), *Varieties of Attention* (s. 515-549). Academic Press.
- Roelofs, A. (2005a). Spoken word planning, comprehending, and self-monitoring: Evaluation of WEAVER++. W: R. J. Hartsuiker, R. Bastiaanse, A. Postma, i F. Wijnen (Red.), *Phonological encoding and monitoring in normal and pathological speech* (s. 42-63). Psychology Press.

- Roelofs, A. (2005b). From Popper to Lakatos: A case for cumulative computational modeling. W: A. Cutler (Red.), *Twenty-first century psycholinguistics: Four cornerstones* (s. 313-330). Routledge.
- Roelofs, A. (2020a). Self-Monitoring in Speaking: In Defense of a Comprehension-Based Account. *Journal of Cognition*, 3(1), 18, 1-13. <https://doi.org/10.5334/joc.61>
- Roelofs, A. (2020b). On (Correctly Representing) Comprehension-Based Monitoring in Speaking: Rejoinder to Nozari (2020). *Journal of Cognition*, 3(1), 20, 1-7. <https://doi.org/10.5334/joc.112>
- Roelofs, A., Piai, V., i Schriefers, H. (2011). Selective attention and distractor frequency in naming performance: Comment on Dhooge and Hartsuiker (2010). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(4), 1032–1038. <https://doi.org/10.1037/a0023328>
- Rosinski, R. R. (1977). Picture-word interference is semantically based. *Child Development*, 48(2), 643–647. <https://doi.org/10.2307/1128667>
- Sabin, E. J., Clemmer, E., O'Connell, D. C., i Kowal, S. (1979). A pausological approach to speech development. W: A. W. Siegman i S. Feldstein (Red.), *Of speech and time: Temporal speech patterns in interpersonal contexts* (s. 35-55). Erlbaum
- Salonen, T., i Laakso, M. (2009). Self-repair of speech by four-year-old Finnish children. *Journal of Child Language*, 36, 855-882. <https://doi.org/10.1017/S0305000908009240>
- Sasisekaran, J., i Weber-Fox, C. (2012). Cross-sectional study of phoneme and rhyme monitoring abilities in children between 7 and 13 years. *Applied Psycholinguistics*, 33, 253–279. <https://doi.org/10.1017/S0142716411000348>
- Schriefers, H., Meyer, A. S., i Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29(1), 86–102. [https://doi.org/10.1016/0749-596X\(90\)90011-N](https://doi.org/10.1016/0749-596X(90)90011-N)
- Searle, J. R. (1980). Czym jest akt mowy? *Pamiętnik Literacki*, 71(2), 241-248.
- Severens, E., Kühn, S., Hartsuiker, R. J., i Brass, M. (2012). Functional mechanisms involved in the internal inhibition of taboo words. *Social Cognitive and Affective Neuroscience*, 7(4), 431-435. <https://doi.org/10.1093/scan/nsr030>
- Seyfeddinipur, M., i Kita, S. (2001). Gesture as an indicator of early error detection in self-monitoring of speech. W: *Proceedings of the ITRW on Disfluency in Spontaneous Speech (DiSS 2001)* (s. 29-32).

- Seyfeddinipur, M., Kita, S., i Indefrey, P. (2008). How speakers interrupt themselves in managing problems in speaking: Evidence from self-repairs. *Cognition*, 108(3), 837-842. <https://doi.org/10.1016/j.cognition.2008.05.004>
- Shao, Z., Roelofs, A., i Meyer, A. S. (2012). Sources of individual differences in the speed of naming objects and actions: The contribution of executive control. *The Quarterly Journal of Experimental Psychology*, 65(10), 1927-1944. <https://doi.org/10.1080/17470218.2012.670252>
- Shao, Z., Meyer, A. S., i Roelofs, A. (2013). Selective and nonselective inhibition of competitors in picture naming. *Memory & Cognition*, 41, 1200–1211. <https://doi.org/10.3758/s13421-013-0332-7>
- Schiller, N. O. (2005). Verbal self-monitoring. W: A. Cutler (Red.), *Twenty-First Century Psycholinguistics: Four Cornerstones* (s. 245-261). Lawrence Erlbaum Associates.
- Schiller, N. O., Jansma-Schmitt, B. M., Peters, J. C., i Levelt, W. J. M. (2006). Monitoring metrical stress in polysyllabic words. *Language and Cognitive Processes*, 21(1-3), 112-140. <https://doi.org/10.1080/01690960400001861>
- Sechrest, L., i Yeaton, W. H. (1982). Magnitudes of experimental effects in social science research. *Evaluation Review*, 6(5), 579-600.
- Sikora, K., Roelofs, A., Hermans, D., i Knoors, H. (2016). Executive control in spoken noun-phrase production: Contributions of updating, inhibiting, and shifting. *The Quarterly Journal of Experimental Psychology*, 69(9), 1719–1740. <https://doi.org/10.1080/17470218.2015.1093007>
- Singer, B. D., i Bashir, A. S. (1999). What are executive functions and self-regulation and what do they have to do with language-learning disorders? *Language, Speech, and Hearing Services in Schools*, 30(3), 265–273. <https://doi.org/10.1044/0161-1461.3003.265>
- Skimina, E., Harasimczuk, J., i Ciecuch, J. (2022). *Podstawowe standardy edytorskie naukowych tekstów psychologicznych w języku polskim na podstawie reguł APA 7*. Wydawnictwo Liberi Libri. <https://doi.org/10.47943/lib.9788363487560>
- Slobin, D. I. (1980). Poznawcze przesłanki rozwoju gramatyki [Reprint]. W: G. W. Shugar i M. Smoczyńska (Red.), *Badania nad rozwojem języka dziecka. Wybór prac*. (s. 398-451). Państwowe Wydawnictwo Naukowe.
- Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C. C., Pechmann, T., Pleh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., Tzeng, O.,

- Arevalo, A., Vargha, A., Butler, A. C., Buffington, R., i Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, 51(2), 247-250. <https://doi.org/10.1016/j.jml.2004.03.002>
- Szepietowska, E.M. (2010). Płynność słowna w neuropsychologii – aspekty teoretyczne i metodologiczne. W: K. Jodzio & E.M. Szepietowska (Red.), *Neuronalne ścieżki poznania i zachowania*. Rozważania interdyscyplinarne (s. 15-28). Wydawnictwo UMCS.
- Tabakowska, E. (1999). Gramatyka i obrazowanie: Wprowadzenie do językoznawstwa kognitywnego. Kraków: Wydawnictwo Oddziału Polskiej Akademii Nauk.
- Tarkowski, Z. (1992). *Jąkanie wczesnodziecięce*. Wydawnictwa Szkolne i Pedagogiczne
- Taylor, J. R. (2007). *Gramatyka kognitywna*. Towarzystwo Autorów i Wydawców Prac Naukowych Universitas.
- Tomasello, M. (2002). *Kulturowe źródła ludzkiego poznawania*. Państwowy Instytut Wydawniczy.
- Tomaszewski, P. (2004). Polski Język Migowy (PJM) – mity i fakty. *Poradnik Językowy*, 6, 59-72.
- Tomczak, M., i Tomczak, E. (2014). The need to report effect size estimates revisited: An overview of some recommended measures of effect size. *Trends in Sport Sciences*, 1(21), 19-25.
- Trudeau-Fisette, P., Ito, T., i Ménard, L. (2019). Auditory and somatosensory interaction in speech perception in children and adults. *Frontiers in Human Neuroscience*, 13, Artykuł 344. <https://doi.org/10.3389/fnhum.2019.00344>
- Van Beinum, F. J. (2008). Frames and Babbling in Hearing and Deaf Infants. W: B. L. Davis i K. Zajdó (Red.), *The Syllable in Speech Production. Perspectives on the Frame Content Theory* (s. 225-241). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203837894>.
- Van den Wildenberg, W. P. M., i Christoffels, I. K. (2010). STOP TALKING! Inhibition of speech is affected by word frequency and dysfunctional impulsivity. *Frontiers in Psychology*, 1, Article 145. <https://doi.org/10.3389/fpsyg.2010.00145>
- Van der Stelt, J. M. (2008). Teething, Chewing, and the Babbled Syllable. W: B. L. Davis i K. Zajdó (Red.), *The Syllable in Speech Production. Perspectives on the Frame Content Theory* (s. 245-272). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203837894>.

- Van Wijk, C., i Kempen, G. (1987). A dual system for producing self-repairs in spontaneous speech: Evidence from experimentally elicited corrections. *Cognitive Psychology*, 19(4), 403–440. [https://doi.org/10.1016/0010-0285\(87\)90014-4](https://doi.org/10.1016/0010-0285(87)90014-4)
- Vasić, N., i Wijnen, F. (2005). Stuttering as a monitoring deficit. W: R. J. Hartsuiker, R. Bastiaanse, A. Postma, i F. Wijnen (Red.), *Phonological Encoding and Monitoring in Normal and Pathological Speech* (s. 226-247). Psychology Press.
- Vigliocco, G., Antonini, T., i Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, 8(4), 314-317
- Vigliocco, G., i Hartsuiker, R. J. (2005). Maximal Input and Feedback in Production and Comprehension. W: A. Cutler (Red.), *Twenty-First Century Psycholinguistics: Four Cornerstones* (s. 209-228). Lawrence Erlbaum Associates.
- Vitkovitch, M., Humphreys, G. W., i Lloyd-Jones, T. J. (1993). On Naming a Giraffe a Zebra: Picture Naming Errors Across Different Object Categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(2), 243-259.
- Wardlow, L. (2013). Individual differences in speaker's perspective taking: The roles of executive control and working memory. *Psychonomic Bulletin & Review*, 20(4), 766-772. <https://doi.org/10.3758/s13423-013-0396-1>
- Wheeldon, L. R., i Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language*, 34(3), 311-334. <https://doi.org/10.1006/jmla.1995.1014>
- Wierzbicka, A. (1999). *Język-Umysł-Kultura*. Wydawnictwo Naukowe PWN.
- Wierzbicka, A. (2011). Uniwersalia ugruntowane empirycznie. *Teksty Drugie. Teoria literatury, krytyka, interpretacja*, (1-2), 13-30.
- Wijnen, F. (1992). Incidental Word and Sound Errors in Young Speakers. *Journal of Memory and Language*, 31, 73–155.
- Wiktorowicz, J., Grzelak, M. M., i Grzeszkiewicz-Radulska, K. (2020). *Analiza statystyczna z IBM SPSS Statistics*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Wolna, A., Łuniewska, M., Haman, E., & Wodniecka, Z. (2023). Polish norms for a set of colored drawings of 168 objects and 146 actions with predictors of naming performance. *Behavior Research Methods*, 55(5), 2706–2732. <https://doi.org/10.3758/s13428-022-01923-3>
- Woźniak, T. (2008). Standard postępowania logopedycznego w przypadku gielkotu. *Logopedia*, 37, 227-23

- Woźniak-Prus, M., Gambin, M. (2011). Metody badania funkcji wykonawczych u dzieci i młodzieży – omówienie wybranych technik. W: M. Świącicka (Red.), *Metody diagnozy w psychologii klinicznej dziecka i rodziny*. Wydawnictwo Paradygmat.
- Wróbel, S. (2007). Modularność w ogniu pytań. W: S. Wróbel (Red.), *Modularność umysłu* (s. 5-9). Wydział Pedagogiczno Artystyczny UAM
- Wygotski, L. S. (1934/1989). *Myślenie i mowa*. Państwowe Wydawnictwo Naukowe.
- Xue, G., Aron, A. R., i Poldrack, R. A. (2008). Common neural substrates for inhibition of spoken and manual responses. *Cerebral Cortex*, 18(8), 1923–1932.  
<https://doi.org/10.1093/cercor/bhm220>
- Zarębina, M. (1965). *Kształtowanie się systemu językowego dziecka*. Zakład Narodowy imienia Ossolińskich.
- Żywiczyński, P., i Wacewicz, S. (2015). *Ewolucja języka. W: stronę hipotez gesturalnych*. Wydawnictwo Naukowe UMK.

**Źródła internetowe (wystąpienia konferencyjne opublikowane w Internecie):**

- Gyarmathy, D., i Neuberger, T. (2013). Self-monitoring strategies: the factor of age. (Travaux du 19eme CIL/19th ICL Papers).  
[https://www.researchgate.net/publication/280077082\\_Self-monitoring\\_strategies\\_the\\_factor\\_of\\_age](https://www.researchgate.net/publication/280077082_Self-monitoring_strategies_the_factor_of_age)
- Ratner, N., i Wijnen, F. (2007). The vicious cycle: Linguistic encoding, self-monitoring and stuttering. (Proceedings of the 5th World Congress of Fluency Disorders. Dublin, Ireland).  
[https://www.researchgate.net/publication/46683234\\_The\\_vicious\\_cycle\\_Linguistic\\_encoding\\_self-monitoring\\_and\\_stuttering](https://www.researchgate.net/publication/46683234_The_vicious_cycle_Linguistic_encoding_self-monitoring_and_stuttering)

## Spis tabel

Tabela 1 Zestawienie porównywanych modeli produkcji mowy	71
Tabela 2 Kamienie milowe w rozwoju fonacji i artykulacji w okresie niemowlęctwa	132
Tabela 3 Zmienne zależne	191
Tabela 4 Zmienne niezależne podmiotowe	193
Tabela 5 Zmienne niezależne manipulacyjne	195
Tabela 6 Charakterystyka uczestników badania w poszczególnych grupach wiekowych	201
Tabela 7 Średnie i odchylenia standardowe procentu zaklasyfikowanych odpowiedzi w zadaniu <i>Nazywania obrazków</i>	207
Tabela 8 Wyrazy (nazwy obrazków) wykorzystane w warunku zgodności ze wskazaniem miejsca artykulacji sylaby docelowej	210
Tabela 9 Wyrazy (nazwy obrazków) wykorzystane w warunku niezgodności	211
Tabela 10 Średnie i odchylenia standardowe procentu zaklasyfikowanych odpowiedzi w zadaniu <i>Monitorowanie sylab</i>	214
Tabela 11 Średnie i odchylenia standardowe procenta odpowiedzi z zadania Powtarzanie zdań zaklasyfikowanych do dalszej analizy	218
Tabela 12 Średnie tempo mówienia w każdej z badanych grup wiekowych w warunku neutralnym zadania <i>Powtarzania zdań</i>	220
Tabela 13 Średnie i odchylenia standardowe dla procenta poprawnych odpowiedzi oraz procenta błędów perseweracyjnych w teście <i>WCST</i> dla trzech grup wiekowych	222
Tabela 14 Średnie i odchylenia standardowe dla procenta poprawnych odpowiedzi w zadaniu <i>1- i 2-wstecz</i>	225
Tabela 15 Statystyki opisowe dla średnich poprawności reakcji w zadaniu z flankerami	227
Tabela 16 Statystyki opisowe poprawności nazywania obrazków	229
Tabela 17 Statystyki opisowe dla średnich czasu nazywania obrazków	230
Tabela 18 Porównanie między grupami średnich czasów <i>Nazywania obrazków</i> w	237

warunku neutralnym	
Tabela 19 Wyniki porównań testem Wilcoxon dla dystrybucji pauz wypełnionych między warunkami zadania Nazywanie obrazków w trzech badanych grupach wiekowych	241
Tabela 20 Statystyki opisowe dla poprawności odpowiedzi w zadaniu <i>Monitorowanie sylab</i>	242
Tabela 21 Statystyki opisowe poprawności odpowiedzi w zadaniu <i>Monitorowanie sylab</i> po transformacji arcsin w trzech grupach wiekowych	242
Tabela 22 Średnie i odchylenia standardowe dla średnich czasów reakcji w warunku zgodności	244
Tabela 23 Średnie i odchylenia standardowe dla średnich czasów reakcji (ms) w warunku niezgodności	244
Tabela 24 Wyniki porównań parami (post hoc) między grupami wiekowymi pod względem poprawności odpowiedzi w zadaniu <i>Monitorowanie sylab</i>	247
Tabela 25 Analiza różnic w średnim czasie reakcji na sylabę docelową w zależności od pozycji w wyrazie i grupy wiekowej	249
Tabela 26 Analiza różnic w czasie monitorowania sylab między wyrazami 2- i 3-sylabowymi w zależności od wieku uczestników	251
Tabela 27 Statystyki opisowe dla długości wypowiedzi reparamum (liczba sylab) w warunku błędnego podmiotu	254
Tabela 28 Statystyki opisowe dla procenta odpowiedzi z reperacją ukrytą (przez odroczenie) w warunku błędnego dopełnienia w trzech grupach wiekowych	254
Tabela 29 Procent odpowiedzi z interupcją w warunku błędnego podmiotu	257
Tabela 30 Statystyki opisowe dla wskaźników poziomu funkcji wykonawczych	267
Tabela 31 Wartość współczynników korelacji między odpornością selekcji leksykalnej na dystrakcję a poziomem funkcji wykonawczych u dzieci	270
Tabela 32 Wartość współczynników korelacji dla analizy związków między efektem interferencji semantycznej a poziomem funkcji wykonawczych	271
Tabela 33 Wartości współczynników korelacji uzyskane w badaniu Shao i in (2012)	272

Tabela 34 Wartość współczynników korelacji między wskaźnikami funkcji wykonawczych a poprawnością oraz szybkością nazywania obrazków w warunku bez interferencji (warunek neutralny)	272
Tabela 35 Wartość współczynników korelacji między sprawnością monitorowania wewnętrznego a poziomem funkcji wykonawczych	274
Tabela 36 Wartość współczynników korelacji między dokonywaniem reperacji a poziomem funkcji wykonawczych	275
Tabela 37 Efekty wpływu zewnętrznej dystrakcji semantycznej na wykonanie zadania <i>Nazywanie obrazków</i> : porównanie trzech grup wiekowych	280
Tabela 38 Zestawienie częstości wystąpień w NKJP wyrazów użytych w warunku zgodności i niezgodności	290

## Spis rysunków

Rysunek 1 Znak językowy wg Hjelmsleva	43
Rysunek 2 Poziomy produkcji mowy według Garreta	54
Rysunek 3 Model produkcji mowy według Levelta (1989)	56
Rysunek 4 Produkcja mowy według Levelta (1999)	66
Rysunek 5 Kanały sprzężenia zwrotnego podczas mówienia	76
Rysunek 6 Przykładowy błąd mowy z reperacją	80
Rysunek 7 Zmiany rozwojowe w procesie produkcji mowy w oparciu o model Levelta	129
Rysunek 8 Wybrane aspekty kontroli procesu mówienia na przykładzie przejęzyczenia i jego reperacji	159
Rysunek 9 Układ zmiennych	197
Rysunek 10 Ilustracja sekwencji i czasu prezentacji bodźców w zadaniu nazywania obrazków	206
Rysunek 11 Ilustracja miejsca artykulacji sylab docelowych w nazwach obrazków	210
Rysunek 12 Ilustracja sekwencji zdarzeń i czasu prezentacji bodźców w przykładowej próbie zadania <i>Monitorowanie sylab</i>	212
Rysunek 13 Ilustracja sekwencji zdarzeń i czasu prezentacji bodźców w przykładowej próbie zadania <i>Powtarzanie zdań</i>	215
Rysunek 14 Przykład spodziewanych reakcji osób badanych w trzech warunkach zadania <i>Powtarzanie zdań</i>	217
Rysunek 15 Fragment ekranu z czterema kartami wzorcowymi i przykładową kartą bodźcową w elektronicznej wersji zadania <i>WCST (BCST)</i>	221
Rysunek 16 Wyniki porównań poprawności nazywania obrazków w warunkach neutralnym między grupami wiekowymi	231
Rysunek 17 Średni procentowy rozkład odpowiedzi udzielanych przez dzieci w warunkach interferencji	234
Rysunek 18 Średnie czasów reakcji w zadaniu <i>Nazywanie obrazków</i>	236
Rysunek 19 Procentowe opóźnienie nazywania w trzech grupach wiekowych wywołane dystrakcją semantyczną	239
Rysunek 20 Wykres normalności K-K z trendem dla rozkładu poprawności odpowiedzi po transformacji arcsin w grupie osób dorosłych	243
Rysunek 21 Wykres K-K z trendem dla rozkładu średnich czasu reakcji osób dorosłych na sylabę docelową w nagłosie	245

Rysunek 22 Wykres K-K z trendem dla rozkładu średnich czasu reakcji osób dorosłych na sylabę docelową w wygłosie	246
Rysunek 23 Czas reakcji w zadaniu <i>Monitorowanie sylab</i> w zależności od pozycji sylaby docelowej w wyrazie: porównanie grup wiekowych	249
Rysunek 24 Czas reakcji w zadaniu <i>Monitorowanie sylab</i> w zależności od długości wyrazu: porównanie grup wiekowych	251
Rysunek 25 Porównanie rozkładów średniej długości wypowiedzi z reparamum w trzech grupach wiekowych	256
Rysunek 26 Rozkłady częstości występowania interupcji w odpowiedziach warunku błędnego podmiotu: porównanie grup wiekowych	258
Rysunek 27 Wykres rozrzutu dla relacji między średnim tempem mówienia, a szybkością dokonywania interupcji	260
Rysunek 28 Proporcje częstości odpowiedzi w warunku błędnego dopełnienia: zestawienie dla trzech grup wiekowych	263
Rysunek 29 Zestawienie różnic w tempie mówienia między zdaniem bez błędu a zdaniem z reperacją ukrytą	266
Rysunek 30 Średnie wartości czasów reakcji we wszystkich pomiarach (oba warunki) zadania <i>Monitorowanie sylab</i>	289

# Załączniki